

Topology for gaze analyses - Raw data segmentation

Oliver Hein and Wolfgang H. Zangemeister
Neurological University Clinic Hamburg UKE

Recent years have witnessed a remarkable growth in the way mathematics, informatics, and computer science can process data. In disciplines such as machine learning, pattern recognition, computer vision, computational neurology, molecular biology, information retrieval, etc., many new methods have been developed to cope with the ever increasing amount and complexity of the data. These new methods offer interesting possibilities for processing, classifying and interpreting eye-tracking data. The present paper exemplifies the application of topological arguments to improve the evaluation of eye-tracking data. The task of classifying raw eye-tracking data into saccades and fixations, with a single, simple as well as intuitive argument, described as coherence of spacetime, is discussed, and the hierarchical ordering of the fixations into dwells is shown. The method, namely identification by topological characteristics (ITop), is parameter-free and needs no pre-processing and post-processing of the raw data. The general and robust topological argument is easy to expand into complex settings of higher visual tasks, making it possible to identify visual strategies.

Keywords: gaze trajectory; event detection; topological data analysis (TDA); clustering; parameter-free classification; visual strategy; global scanpath; local scanpath

Introduction

Gaze trajectories can tell us many interesting things about human nature, including attention, memory, consciousness, etc., with important applications (Groner & Groner, 1982; Duchowski, 2002; Van der Stigchel, Meeter, & Theeuwes, 2006; Russo, 2010) as well as facilitating the diagnosis and helping to understand the mechanisms of diseases (Leigh & Kennard, 2004; Munoz, Armstrong, & Coe, 2007; Crabb et al., 2010). Normally, viewing behavior is studied with simple paradigms to keep the complexity of natural viewing situations as low as possible, e.g., in a search paradigm, a person looks at a computer screen with a simple static geometric configuration under well defined optical constraints, i.e., constant illumination, head immobilized by a chin rest or bite bar, no distractors, etc.

The task of analyzing, classifying, and interpreting gaze trajectories for realistic situations proves to be

much more difficult because of the many different factors influencing the steering of the eyes. The usual scientific approach is to break down real world complexity into easy to define and control partial modules, and then to try to reassemble reality from these simple modules. This has also been done for gaze trajectories. The task of analyzing the gaze trajectory data can roughly be split into two subtasks: the low level description of the noisy raw data that are produced from the gaze tracker, and the high level description of the data in combination with the viewing task and the cognitive processes. The first subtask could be regarded as the mathematical modeling of high frequency time-series, given that modern gaze trackers can sample eye position and orientation at 2000 Hz or even more (Andersson, Nyström, & Holmqvist, 2010).

The careful choice of the data model and data representation is the basis for all of the following analyses. Only a model capable of incorporating the many subtleties of the gaze trajectory is able to support the complex questions which appear in the context of modeling the looking task in relation to the assumed cognitive processes.¹ Of course, a more complex model is harder to implement and interpret. There is a permanent balancing between data load, explanatory potential, and

¹Realistically the model is a strong assumption (prior) and very often the hypothesized construct is driven by the original model.

History: Received July 24, 2016; Published March 13, 2017.
Citation: Hein, O. & Zangemeister, W. H. (2017). Topology for gaze analyses - Raw data segmentation. *Journal of Eye Movement Research*, 10(1):1, 1-25.
Digital Object Identifier: 10.16910/jemr.10.1.1
ISSN: 1995-8692
This article is licensed under a Creative Commons Attribution 4.0 International license 

model complexity.

Splitting trajectory data into events

In this section a general outline of splitting raw eye-tracking data into meaningful events is given. At present, the most important segmentation of the data is the dichotomous splitting into fixations and saccades. Although this is a long standing approach, up to now no definite algorithm for the splitting exists. The reasons are discussed.

The basic oculomotor events

The eyes' scanning of the surrounding is done in a sequential manner, since the movement of the eyes, seen as a mechanical system, is limited to sequential movements. It has to be remarked that, in many aspects, this is not true for the information extraction and processing of the visual data within the brain, which can process information in parallel (Thornton & Gilden, 2007; Trukenbrod & Engbert, 2012). It is well known that a detailed analysis can only be done for a very small part of the visual scene, approximately 1 up to 5 degrees of visual angle (Carpenter, 1988; Duchowski, 2007). This is the part of the scene which is projected onto the fovea, the region of the retina with the highest concentration of cone cells. To capture the whole scene, the eyes have to switch swiftly to other regions within the scene, which is done via saccades, i.e., very fast movements (Gilchrist, 2011; Land, 2011). In fact, saccades are operationally defined by velocity, acceleration, and amplitude criteria. Saccades exhibit a clear characteristic, which is relatively stable across subjects (Leigh & Zee, 2006). Quantitatively this relationship is expressed in the main sequence (Bahill, Clark, & Stark, 1975; Bahill, Brockenbrough, & Troost, 1981; Bahill, 1983; Bollen et al., 1993). Speed is crucial, because the brain has to integrate many parts of the whole scene into one consistent and stable internal representation of our surrounding world, and because of the fact that the observer has decreased sensitivity while the eyes are moving fast, a phenomenon called saccadic suppression (Matin, 1974; Leigh & Zee, 2006). Information gathering works by swiftly scanning the scene and minimizing the timespan of decreased sensitivity. This fact makes a bipartition of the gaze trajectory data desirable.

The gaze trajectory is broken down into two general subsegments, fixations and saccades. Saccades allow the gaze to change between parts of the scene, while fixations are intended for analyzing parts of the scene. Saccades are the segments of the trajectory where eyes are moving fast and in a preprogrammed, directed

manner, whereas in a fixation eyes are moving slowly and in a random-like fashion (Rolfs, 2009). The two modes of movement are displayed alternatively and exclusively. Fixations may then be defined as the part between the saccades or vice-versa. This is a sensible and convenient assumption, but also a major simplification. It is well known that fixations can contain microsaccades as subitems (Martinez-Conde, Macknik, Troncoso, & Hubel, 2009; Rolfs, 2009; Engbert, Mergenthaler, Sinn, & Pikovsky, 2011), mixing the two assumed modes of movement.

These two different movement characteristics can be operationalized. The bipartite classification of gaze points in saccade points and fixation points is normally achieved through a combination of space and time characteristics, i.e., for a fixation, the dispersion of the gaze points on the display combined with the duration of a cluster of gaze points in time; for a saccade, it is the velocity, acceleration, and amplitude of the movement. The exact determination of the parameters and the algorithmic implementation has a long history and many parameterizations exist.²

The classification of eye movements into fixations and saccades is by no means straightforward. One always has to bear in mind that the dichotomic splitting of the data follows our desire for simple and parsimonious models,³ it is not Nature's design. It has to be noted that the eye has a much broader repertoire of movements (Liversedge, Gilchrist, & Everling, 2011). "Patterns" of eye movements other than fixations and saccades occur in real data, e.g., vestibular and vergence eye movements, dynamic over-/undershooting, microsaccades, drift, tremor, etc. This becomes even more complex when viewing dynamic scenes as opposed to still images (Crabb et al., 2010). Because of the moving content, the eyes have to follow the in-

² (Mason, 1976; Karsh & Breitenbach, 1983; Widdel, 1984; Scinto & Barnette, 1986; Stampe, 1993; Krauzlis & Miles, 1996; Wyatt, 1998; Salvucci & Goldberg, 2000; Privitera & Stark, 2000; Larsson, 2002; Engbert & Kliegl, 2003; Smeets & Hooge, 2003; Santella & DeCarlo, 2004; Engbert & Mergenthaler, 2006; Urruty, Lew, & Ihadaddene, 2007; Špakov & Miniotas, 2007; Shic, Scassellati, & Chawarska, 2008; Camilli, Nacchia, Terenzi, & Nocera, 2008; Kumar, Klingner, Puranik, Winograd, & Paepcke, 2008; Munn, Stefano, & Pelz, 2008; Blignaut, 2009; Komogortsev, Jayarathna, Koh, & Gowda, 2009; Nyström & Holmqvist, 2010; Komogortsev, Gobert, Jayarathna, Koh, & Gowda, 2010; Dorr, Jarodzka, & Barth, 2010; van der Lans, Wedel, & Pieters, 2011; Mould, Foster, Amano, & Oakley, 2012; Komogortsev & Karpov, 2012/13; Vidal, Bulling, & Gellersen, 2012; Liston, Krukowski, & Stone, 2012; Špakov, 2012; Valsecchi, Gegenfurtner, & Schütz, 2013)

³Entia non sunt multiplicanda praeter necessitatem (Entities must not be multiplied beyond necessity). –John Punch-

focus part of the scene. The concept of a fixation as being localized in a small subregion of a still image is no longer valid and has to be replaced by the concept of smooth pursuit (Blackmon, Ho, Chernyak, Azzariti, & Stark, 1999). As of now the most important event types are fixations, saccades and smooth pursuit. More recently post-saccadic oscillations (PSO) have come into focus (Nyström & Holmqvist, 2010; Andersson, Larsson, Holmqvist, Stridh, & Nyström, 2016). Zemblys, Niehorster, Komogortsev, and Holmqvist (2017) estimate 15-25 events that, as of now, have been described in the psychological and neurological eye-movement literature.

As common for biological systems, all movements exhibit a normal physiological variability (Smeets & Hooge, 2003; van der Lans et al., 2011). Different application regimes also show different characteristics, e.g., normal reading is different from reading a drifting text (Valsecchi et al., 2013) as it is now common when reading, or even browsing, texts on mobile devices (swiping the text). Furthermore, gaze tracking data can be interrupted by blinks. Blinks interrupt the flow of gaze tracking data, while the eye is still moving consistently. Though coupled (Horn & Adamczyk, 2012), blinks are considered noise.

Even if all possible events were known and clearly defined, the algorithmic processing would introduce a bias into the results. There are many reasons for this finding. One reason lies in the different sensitivities to noise and filter effects (Inchingolo & Spanio, 1985; Tole & Young, 1981), e.g., numerical differentiation is an operation with notorious "bad behavior". Furthermore, the filters used for preprocessing also call for parameters and introduce a bias into the data.

Higher level use for oculomotor events

Another motivation for the development of more and more sophisticated algorithms is the growing – one might say exploding – applicability of eye tracking devices. In the past eye tracking was restricted to scientific uses and the tasks people were performing were relatively low in complexity, e.g., a simple search task. Nowadays, with the increase of performance in eye-tracking hardware and computing power, the tasks under investigation have become more and more complex, producing a wealth of data.

Recent years especially have shown a growing interest in the investigation of complex dynamic settings. In these settings the viewing subject is no longer looking at a static image from a (head-)fixed position. In the extreme, the subject is moving freely and interacting with its environment, like playing table tennis or driving a car (Land & Lee, 1994; Land & Furneaux, 1997;

Land & Tatler, 2009; Lappi & Lehtonen, 2013). Driven by industrial applications such as market research, dynamic scenes are playing a more and more important role. These can be watching TV and movies (Goldstein, Woods, & Peli, 2007; Brasel & Gips, 2008; Dorr, Vig, & Barth, 2012), video clips (Carmi & Itti, 2006; Berg, Boehnke, Marino, Munoz, & Itti, 2009; Tseng, Carmi, Cameron, Munoz, & Itti, 2009) or interactively playing a video game (Peters & Itti, 2008; Sundstedt, Stavrakis, Wimmer, & Reinhard, 2008). Another application is the assessment of the driving ability in diseases like glaucoma (Crabb et al., 2010) or Parkinson's Disease (Buhmann et al., 2014), where patients view hazardous situations in a car driving context. The system calibration can be automated, allowing the collection of data for many subjects. As an example, the eye movements of 5,638 subjects have successfully been recorded while they viewed digitized images of paintings from the National Gallery collection in the course of the millennium exhibition (Wooding, 2002b; Wooding, Muggleston, Purdy, & Gale, 2002; Wooding, 2002a). It is apparent that such data sets can not be evaluated manually. A recent application is online tracking of eye movements for integration in gaze contingent applications, e.g., driving assistance, virtual reality, gaming, etc. Here the online tracking produces a continuous stream of highly noisy data, and the system has to extract the relevant events in real time and has to infer the users' intents to adjust itself to their needs.

These more complex settings and large sample sizes are not only a challenge for the hard- and software, but also require a rethinking of the concepts being used to interpret the data, especially when it comes to the theoretical possibility of inferring people's intent from their eye movements (Haji-Abolhassani & Clark, 2014; Borji & Itti, 2014; Greene, Liu, & Wolfe, 2012).

In summary, the analysis of eye tracking data can be organized in a hierarchy spanning different scales, going from low level segmentation ascending to higher levels, relevant for the physiological and psychological interpretation. Topmost is the comparison and analysis of different eye movement patterns within and between groups of people, as is relevant for the inference of underlying physiological and cognitive processes, which forms the basis for important eye tracking applications, see table 1. Highlighted in light gray background is the first level aggregation into basic events. Highlighted in dark gray is the second level aggregation for higher use, i.e., sets of sequential fixations in a confined part of the viewing area (Santella & DeCarlo, 2004).⁴

⁴In reading called gaze (Just & Carpenter, 1980) and in hu-

Table 1
Functional overview

Module	Description
Preprocessing	<ul style="list-style-type: none"> Noise reduction, smoothing Outlier and blink elimination Gap filling
First level segmentation	<ul style="list-style-type: none"> Fixation, saccade, smooth pursuit, post-saccadic oscillation, ... Fixation statistics, BCEA,⁵... Main sequence, saccade velocity profile, ...
Postprocessing	<ul style="list-style-type: none"> Removal of physiologically implausible short fixations and saccades Merging of nearby fixations Human assessment
Higher level segmentation	<ul style="list-style-type: none"> Scan path representation Dwell, glance, gaze, ... Area (Region) of interest Viewing strategy
Process inference	<ul style="list-style-type: none"> Comparison within subject and between subjects Modeling, e.g., bottom-up saliency, stochastic processes, ...
Application	<ul style="list-style-type: none"> Medicine, Biometrics, VR, Gaming ...

The problem of defining a fixation

For most areas of inquiry this level of information in the raw data is not necessary. It is sufficient to reduce the gaze-points into oculomotor events, i.e., into the fixations and saccades forming the scanpath. Here scanpath⁶ means any higher level time ordered representation of the raw data which form the physical gaze trajectory. The fixations can further be attributed to regions of interest (RoI), each RoI representing a larger part of the scene with interesting content for the viewing subject.

While intuitively easy to grasp, it is by no means obvious how to explicitly define these concepts and make them available for numerical calculations (Andersson et al., 2016). Very often only basic saccade and fixation identification algorithms are part of the eye-tracking system at delivery (B. W. Tatler, Wade, Kwan, Findlay, & Velichkovsky, 2010), leaving the higher splitting up to the user. This is desirable in the academic setting, but not in the industrial setting, where time efficient analysis has to be conducted, e.g., in marketing research (Reutskaja, Nagel, Camerer, & Rangel, 2011) or

in usability evaluation (Goldberg & Wichansky, 2003). Most commercial implementations incorporate dispersion threshold methods, e.g., ASL (2007) or velocity threshold methods, e.g., seeingmachines (2005); Olsen (2012); Tobii (2014). Some offer the user flexibility in choosing the thresholds, while others mask the complexity from the user by assuming a sort of lowest common denominator for the thresholds in different application domains, although it is known that parameters can vary between different tasks, e.g., the mean fixation duration amounts to 225 ms on silent reading, 275 ms on visual search, and 400 ms on hand-eye coordination (Rayner, 1998). To account for these variations, some implementations have 10 parameters to adjust (Reimer & Sodhi, 2006), requiring a good understanding of the theory of gaze trajectories.

It is well known that the parametrization of the algorithm can substantially affect the results, but there is no rule which algorithm and which parametrization to employ in a given experimental setting (Smyrnis, 2008; Nyström & Holmqvist, 2010; Wass, Smith, & Johnson, 2013). A comparison of the different algorithms and the bias which can result under different parameterizations is given in Shic et al. (2008); Špakov (2012); Andersson et al. (2016). For instance, post-saccadic oscillations (PSOs), i.e., wobbling over/under-shootings, are usually not explicitly mentioned, but form a normal part of eye movements. The PSOs are attributed to fixations or saccades, influencing the overall statistics of the measurement (Nyström & Holmqvist, 2010; Andersson et al., 2016). The algorithms to implement the classification are therefore different and researchers aim to improve and extend the algorithms constantly (van der Lans et al., 2011; Špakov, 2012; Komogortsev & Karpov, 2012/13; Mould et al., 2012; Liston et al., 2012; Vidal et al., 2012; Wass et al., 2013; Valsecchi et al., 2013; Daye & Optican, 2014; Andersson et al., 2016; Hessels, Niehorster, Kemner, & Hooge, 2016; Zemblyns et al., 2017).

man factors called glance (Green, 2002). To avoid confusion with the standard meaning of gaze the term dwell is used (Holmqvist et al., 2011).

⁵Bivariate Contour Ellipse Area

⁶The term scanpath is somewhat vague and differs in its meaning and interpretation between different research areas and authors. Introduced in 1971 by Noton and Stark (Noton & Stark, 1971b, 1971a, 1971c; Zangemeister, Stiehl, & Freksa, 1996), it was a fairly abstract concept to describe a repetitive pattern of a single subject while viewing a static stimulus (Privitera, 2006). Common terminology has been improved with works such as Holmqvist et al. (2011), research networks such as COGAIN, or industry driven demands such as the ISO 15007 and SAE J2396 standards for in-vehicle visual demand measurements (Green, 2002).

Many researchers agree that a normative definition and protocol is desirable but at present far from becoming reality (Karsh & Breitenbach, 1983; Komogortsev et al., 2010; Nyström & Holmqvist, 2010; Andersson et al., 2016). As Karsh and Breitenbach (1983) stated rightly:

The problem of defining a fixation is one that perhaps deserves more recognition than it had in the past. Generally speaking, the more complex the system the more complex the task of definition will be. ... Once these needs are recognized and implemented, comparison between studies take on considerably more meaning.

Topological approach to the problem

Up until now, no single algorithm has been able to cover all the various aspects in eye tracking data (Andersson et al., 2016). The aim here is to show that there exists a strikingly simple argument for demarcating the different components of the gaze trajectory in a normative way. From well-known approaches a data representation is derived, which forms the basis for a consistent analysis scheme to cover the basic aggregation steps, see gray parts of table 1. The argument for the segmentation is a topological one and is by its very nature global and scale-invariant. It is the mathematical formulation that a fixation is a coherent part in space and time. The meaning of “coherent in space and time” will be clarified in the next sections. The argument needs no thresholds or calibration and is independent of any experimental setting or paradigm. The delineation of the gaze trajectory is unambiguously reproducible.

Overview of existing approaches

This section presents an overview of different approaches to event detection. From these, a common argument is isolated, the coherence of sample data in space and time, which in turn forms the basis for the new algorithm.

Taxonomy of algorithms

At present, we see a wide variety of different methods being used to extract the main oculomotor events from raw eye tracking data (Holmqvist et al., 2011). Each approach to the data highlights at least one prominent and distinguishing feature of the main oculomotor events in the trajectory data and makes use of specialized algorithms to filter/detect these features against the noisy background. Noise is to be understood as being the part of the measurement which is not relevant

for the investigation, e.g., micro saccades can be considered noise in one study, but be of central interest in another setting. In its narrow sense noise is the random part inherent in any measurement. There is a common logic to all these approaches, from which a data representation and global topological argument can be derived. To better understand the topological approach, algorithms currently in use are systematized in a taxonomy. The taxonomy was first introduced in Salvucci and Goldberg (2000). This classification has often been repeated and adapted in the literature (Komogortsev et al., 2010; Koh, Gowda, & Komogortsev, 2010; Komogortsev & Karpov, 2012/13; Santini, Fuhl, Kübler, & Kasneci, 2016; Andersson et al., 2016). Here, as in Salvucci and Goldberg (2000), the classification is based on the role of time and space as well as algorithms used to evaluate raw data. Broadly speaking, there are two different approaches to the data, which differ in complexity.

The algorithmically simplest approach is based on thresholds for saccades and fixations. In the case of saccades these are thresholds for velocity (I-VT: identification by velocity threshold), acceleration, and even jerk, very often calculated as the discrete numerical space-time n -point difference approximations to the continuous differentials. E.g., a saccade is detected whenever the eye’s angular velocity is greater than 30 deg/s (Poulton, 1962; Stampe, 1993; Fischer, Biscaldi, & Otto, 1993; Gitelman, 2002; Paulsen, Hallquist, Geier, & Luna, 2015). These algorithms are called “saccade pickers” (Karn, 2000).

The second group targets the space dispersion (I-DT: identification by dispersion (position-variance) threshold) or space-time dispersion (I-DDT: identification by dispersion and duration thresholds), i.e., when a consecutive series of gaze points occur near each other in display space, they are considered part of a fixation. E.g., in a reading context, a fixation lasts between 200 and 300 msec and a saccade spans approximately seven character spaces (Rayner, 1998). Gaze points consistent with this are aggregated and assumed to form a single fixation. These algorithms are called “fixation pickers”.

Most algorithms use simple thresholds to cluster data into saccades and fixations, which in practice need to be optimized. A fixed parameter approach may perform well on a specific record but is very often too imprecise and error-prone when applied to different records.⁷ In order to improve results, researchers adapt

⁷Two-state Hidden Markov models (HMM) are intrinsically based on fitting individual data thus avoiding the problem of setting parameters explicitly (Salvucci & Goldberg, 2000; Rothkopf & Pelz, 2004). A prerequisite is, however, to assume two states, i.e., saccade and fixation, limiting the

the threshold in a dynamic way (Engbert & Mergenthaler, 2006; Nyström & Holmqvist, 2010), or combine criteria, e.g., a saccade is detected when the angular velocity is higher than 30 deg/s, the angular acceleration exceeds 8000 deg/s², the deflection in eye position is of at least 0.1 deg, and a minimum duration of 4 ms is exceeded (B. Tatler, Wade, & Kaulard, 2007; Frey, Honey, & König, 2008; N. D. Smith, Crabb, Glen, Burton, & Garway-Heath, 2012; T. J. Smith & Mital, 2013). Note that dispersion thresholds can be inversely defined for saccades, i.e., in relations to a fixation, a saccade is over-dispersed, i.e., it has a minimum jumping distance. This is essential when delineating micro saccades from saccades.

Parameters are often chosen subject to individual judgment or even rather arbitrarily (Itti, 2005). Even after using more criteria, human post-processing is required (Wass et al., 2013), and means to reduce the human interaction are being sought (de Bruin, Malan, & Eloff, 2013).

A higher sampling rate of the eye-tracker will give better approximations of velocity and acceleration, but the devices are more expensive and demand higher restrictions for the tested subjects, e.g., a chin rest, etc. It is remarkable that functional relationships like the main sequence (Bahill et al., 1975) are rarely employed, considering that they give good guidance for setting parameter thresholds (Inchingolo & Spanio, 1985); a recent exception is Liston et al. (2012).

All these approaches are purely operational, call for experience, and are driven by technical as well as programming restrictions. More complex algorithms are of course harder to code and often suffer from performance issues. The simple velocity and dispersion based classifiers are exemplified in table 2 (citations contain an explicit exposure of algorithm).

A considerable advantage of these approaches is that thresholds are easy to understand, interpret, and implement. The values for thresholds depend on research domain, e.g., the space-time dispersion values in I-DDT are different in reading and in visual search. Fixation times are domain specific, i.e., the duration of a typical fixation in reading is different to fixation times in visual search, etc. (Rayner, 1998). Hand-tuning is often requisite to get good results and is based on heuristics.

Range of advanced methods

The more sophisticated algorithms use ramified versions of the basic velocity/dispersion features taken from signal processing, statistics, Kalman filtering, Bayesian state estimation, clustering, pattern classifier algorithms, and machine learning.

Table 2
Taxonomy of algorithms

saccade pickers
$\frac{d}{dt}$ velocity threshold I-VT
→ fix (Stampe, 1993)
→ adaptive (Nyström & Holmqvist, 2010)
$\frac{d^2}{dt^2}$ acceleration threshold I-AT
→ fix (Behrens & Weiss, 1992; Behrens, MacKeben, & Schröder-Preikschat, 2010)
$\frac{d^3}{dt^3}$ jerk threshold I-JT
→ fix (Wyatt, 1998), (Matsuoka & Harato, 1983, in Japanese)
fixation pickers
dispersion threshold I-DT
→ fix (Mason, 1976; Kliegl & Olson, 1981)
dispersion and duration thresholds I-DDT
→ fix (Widdel, 1984; Nodine, Kundel, Toto, & Krupinski, 1992; Manor & Gordon, 2003; Krasanakis, Filippakopoulou, & Nakos, 2014)
These are taken from other disciplines like
• Signal processing
– Finite impulse response filter (Tole & Young, 1981)
– Cumulative sum (CUSUM) (Olsson, 2007; Tobii, 2014; Gustafsson, 2000)
• Statistics
– F-test and correlation (Veneri et al., 2010, 2011; Veneri, 2013)
– Gap-statistics (Mould et al., 2012)
• Stochastic processes and time series analysis
– Auto-regressive processes and wavelet analysis (Duchowski, 1998)
• Bayesian approaches
– Hidden Markov model (Salvucci & Anderson, 1998; Rothkopf & Pelz, 2004)
– Kalman filter (Sauter, Martin, Di Renzo, & Vomscheid, 1991; Komogortsev & Khan, 2007)
– Bayesian mixture model (Tafaj, Kasneci, Rosenstiel, & Bogdan, 2012; Kasneci, Kasneci, Kübler, & Rosenstiel, 2014)
– Particle filter (Daye & Optican, 2014)
• Data clustering

classification.

- k-means clustering (Privitera & Stark, 2000)
- Projection clustering (Urruty et al., 2007)
- Mean shift clustering (Santella & DeCarlo, 2004)
- Mean shift clustering and entropy (Vella, Infantino, & Scardino, 2016)
- Two-means clustering (Hessels et al., 2016)
- Machine learning
 - Random forest classifier (Zembly et al., 2017)
 - Neural networks (Hoppe & Bulling, 2016; Anantrasirichai, Gilchrist, & Bull, 2016)
- Graph theory
 - Minimum spanning tree (Goldberg & Schryver, 1995; Rigas, Economou, & Fotopoulos, 2012)
- Fuzzy-set methods
 - (Arzi & Magnin, 1989; Czabanski, Pander, & Przybyla, 2014)
- Shape features⁸
 - Single feature (simple) (Jüttner & Wolf, 1992), (Berg et al., 2009)
 - Multiple features (complex) (Vidal et al., 2012)
 - Mathematical morphology (Longbotham et al., 1994)
- Speech recognition
 - Mel-frequency cepstral analysis (Cuong, Dinh, & Ho, 2012)
- Template matching
 - Velocity-Duration template (Liston et al., 2012)
- Dynamic system analysis
 - Time-delay reconstruction (Shelhamer, 1998; Shelhamer & Zalewski, 2001)

As of now threshold based methods are common standard. Probabilistic methods are promising candidates inasmuch as they offer the possibility to implement an online learning algorithm to adjust to changing viewing behavior. Very recent candidates for event classification are neural networks (Hoppe & Bulling, 2016; Anantrasirichai et al., 2016), random forests (Zembly et al., 2017) or machine learning in general (Zembly, 2016).

Topological data analysis

A relative recent field of data analysis is topological data analysis (TDA). In this section, a topological approach to the data is given. To this end, the notion of different spaces, projections and metrics for the trajectory is introduced. The idea of trajectory spacetime co-

herence is given a precise meaning in topological terms, i.e., “no holes in trajectory spacetime”, a strikingly simple topological argument for the separation of the sample data. An intuition and first use for the argument is given by the visual assessment of the trajectory spacetime, showing the coarse/fine (global/local) structure of a scanpath.

Configuration in physical space

The crucial aspect for partitioning the data is the representation of space and time. Space is here understood as the three-dimensional physical space, called *world space*, which contains as objects the viewer, items viewed, and tracking equipment. Essentially, the viewer’s head and eyes have position (location) and orientation, together called pose, in world space. In the case of the eyes, very often only the direction is determined. The starting point for analysis is the set of raw data from the gaze tracker. The logging of continuous movement of head and eyes consists of the discretely sampled position and orientation of head and eyes in three-dimensional space at equidistant moments in time during the timespan of the experiment.

If it were the intention only to detect fixations or saccades, it would be sufficient to analyze the movement of the eyes in head space. In the context of, e.g., cognitive studies, position and orientation of head and eyes is not interesting in itself; of interest are the visual field, the objects within the visual field and the distribution of allocated attention within the viewer’s internal representation of the visual field, “the objects looked at”. Because of this, the motion of the visual field in world space will be modeled.

The visual field encompasses the part of the environment which is in principle accessible for gathering optical information. It is well known in visual optics that the way of light from an object onto the retina is a multi-stage process which depends on the optical conditions in world space as well as the geometry and refractive power of the different parts of the individual eye (Artal, 2014; Mosquera, Verma, & McAlinden, 2015). Taken together, this is a complex setting to analyze.

In order to cope with the complexity, several assumptions and simplifications have to be made in the course of modeling. The visual field is not directly accessible to the eye tracker. The eye tracker can only measure related signals. These signals are linked by calibration to the point of regard. E.g., in video based head-eye tracking, camera(s) take pictures of the head

⁸E.g., a saccade epoch in the trajectory is in first approximation a straight line, which is a geometric shape feature (a closer look shows it is curved (Inhoff, Seymour, Schad, & Greenberg, 2010)).

and eyes of a subject. The individual images are processed to identify predefined external features of the head and the eyes, e.g., the corners of the mouth and the eyes, the pupil, and glints from light emitting diodes on the light reflecting surfaces of the eyes. From the relative position of these features in image space(s) and the calibration, the gaze⁹ can be determined.

The visual field for one eye is approximated as a right circular cone of one sheet with the gaze-ray as its axis, the center of the entrance pupil as its apex, and with a varying aperture, neglecting any asymmetry of the visual field. For foveated objects the cone angle of a bundle of rays that come to a focus is very small, approximately 0.5 degrees. In the limit of 0.0 degrees only a ray remains, which is convenient for calculations. One calculates the point of intersection of the gaze-ray (starting from the center of the entrance pupil) with an object in world space, and not the projection of the content of the gaze cone onto the retina. Very often one does not work with the gaze-rays of the two eyes separately but instead with only one of the two (the dominant eye); alternatively, the two gaze-rays are combined into a single gaze-ray, i.e., a mean gaze-ray known as “cyclops view” (Elbaum, Wagner, & Botzer, 2017). In addition, very often the head is fixed to prevent head movements at the cost of a somewhat non-physiological setting.

To describe the geometric and topological approach to the data in detail, we will choose the situation where a subject is looking at a screen presenting a visual task (which is a common experimental setting). The point of regard (PoR) is the location toward which the eyes are pointed at a moment in time, i.e., the point of intersection of the (mean) gaze-ray with the screen. Please note that the topological method can work just as well in a three-dimensional setting, e.g., navigating in outdoor scenes. The 3D case is of recent interest for orientation in real and virtual space. For the sake of clarity of explanation, we will now discuss a typical two dimensional setting.

Coherence in space and time

The rationale behind the intended clustering is that trajectory points which have a certain **coherence in space and time** should be grouped together. The question is how to define and express **spacetime coherence** for trajectory points. The argumentation starts with the continuous gaze trajectory tr . The gaze trajectory consists of the time-ordered points of intersection P_i s of the mean gaze-ray with the screen or screen space Σ , within the timespan ts of the experiment. In mathematical abstraction:

$$\mathbb{R} \supset ts \xrightarrow{tr} \Sigma \subset \mathbb{R}^2, tr \subset \mathbb{R} \times \mathbb{R}^2$$

The terminology and notation is not a mathematical pedantism. In the following, different spaces will be introduced and it is essential not to lose track of one’s current conceptual location. It is important to note that the unparametrized P s form a multiset because the gaze-ray can visit the same screen point at many time points (within a fixation and recurrently). Contrary to screen points, a time point, representing an instant or moment in the flow of time, can be visited or passed only once. In practical terms we only have a finite number of discrete data, i.e., the protocol pr of sampled tr . The pr results from a discretization of continuous space and time. The screen consists of a finite number of square pixels all with equal side length $\Delta x = \Delta y = \text{constant}$, the constituting discrete elements of screen space $\Sigma' = \{P_{x,y} : x \in \{0, 1, \dots, 1023\}, y \in \{0, 1, \dots, 767\}\}$, here XGA resolution is assumed, and the tracker takes pictures at moments in time with a constant sampling rate (time points or moments) $ts' = \{M_i : i \in \{0, 1, \dots, N - 1\}\}$, therefore $pr = \{P_{M_0}, P_{M_1}, P_{M_2}, \dots, P_{M_n}\}$. Time is considered to be an ordering parameter, and because of the constant sampling rate, only time index is noted $pr = (P_0, P_1, P_2, \dots, P_n)$ with the ordering parameter $i \in \mathbb{N}_0$. It is important to note that the points of intersection alone do not carry any time information. If we want to convey the information about time ordering, we must label points, i.e., show the index. Graphically we can also show a polyline with the line segments sensed, i.e., showing an arrowhead, see fig. 1.

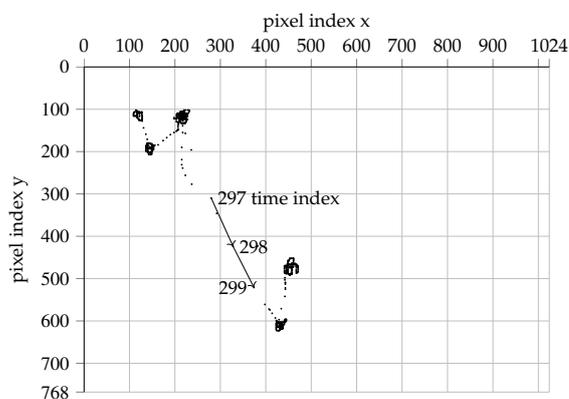


Figure 1. Trajectory in screen space

The crucial step for the following is to take a different position with regard to the subject, the combinatorial view. In analogy to space dispersion algorithms, the

⁹Here gaze is understood as the ray from the center of the entrance pupil and the point-of-regard, essentially the first part of the line-of-sight. For a detailed discussion of the related notions line-of-sight, pupillary axis, visual axis, etc. see (Bennett & Rabbetts, 2007; Schwartz, 2013).

spatial distance of two points is taken, but this time not only for consecutive points in time but all possible 2-point combinations over time. This could be regarded as taking the maximal window size in the dispersion algorithms. This way one obtains the time indexed matrix \mathbf{D} of all combinatorial 2-point distances for the trajectory space. \mathbf{D} serves as the basis for further evaluation. The representation as a time indexed matrix of combinatorial 2-point distances makes the trajectory independent of Euclidean motions because distances are the invariants of Euclidean geometry. The property of being independent of Euclidean motions is especially desirable when comparing scanpaths (Jarodzka, Holmqvist, & Nyström, 2010). At first sight this approach may seem to resemble a superfluous brute force dispersion approach. The advantage of such an approach will be clear from the subsequent sections.

First, we can make the spatio-temporal relationship of the P_i s directly visible with an imaging technique. To this end, we convert, for all time ordered pairs of trajectory points (P_i, P_j) , the screen space distance values $d_{i,j}$ into gray values of a picture, $img(\mathbf{D})$, of size $|pr| \times |pr|$. E.g., when the gaze tracker takes 633 samples one obtains an image measuring 633 by 633 pixels.¹⁰

In the first line fig. 2 should seem suggestive. For the visual system of the human observer, the square block structure of $img(\mathbf{D})$ along the diagonal is easy to identify. The squares along the diagonal represent the fixations. While fixations are spatially confined, their sample distances are short and their gray level is near black. The duration of a fixation is the diagonal (side) length of the square. The first off-diagonal rectangles represent the saccades between successive fixations. Spatially wider saccadic jumps are brighter and shorter jumps are darker. The building blocks form a hierarchy. First level squares are the fixations, second level squares are clusters of fixations, and so on, see fig. 3 (a). The hierarchy of squares along the diagonal is the visual representation for the trajectory (screen)spacetime coherence over different time spans, i.e., the scaling property in time. The scale runs from the base-scale, set by the sampling rate of the tracker, into its first physiological scale, i.e., the time-scale in a single fixation, showing, e.g., tremor, drift, and microsaccades, into the time-scale of several fixations within a dwell, viewing interesting regions, and finally into the time-scale of shifts in interest, changing the viewing behavior.

Visual assessment of trajectory spacetime

The higher level splitting of the viewing behavior in space and time is a much debated subject (Velichkovsky, Joos, Helmert, & Pannash, 2005). The rationale comes under various names in different con-

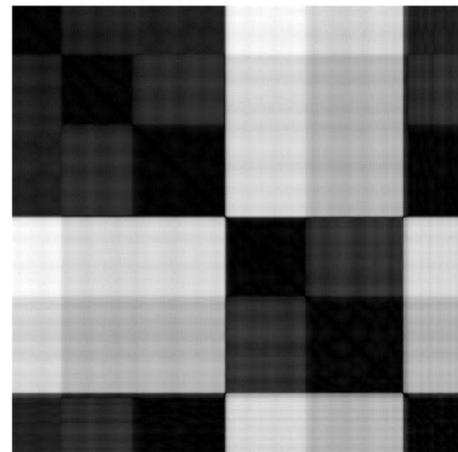


Figure 2. Image of time indexed matrix of 2-point combinatorial distances $img(\mathbf{D})$

texts. At its base, there is a dichotomy in terms of global/local (Groner, Walder, & Groner, 1984; Menz & Groner, 1985; Groner & Groner, 1989), coarse/fine (Over, Hooze, Vlaskamp, & Erkelens, 2007; Godwin, Reichle, & Menneer, 2014), ambient/focal (Helo, Rämä, Pannash, & Meary, 2016), where/what (Sheth & Young, 2016), examining/noticing (Weiskrantz, 1972), which is backed by anatomical findings, i.e., the concept of a ventral and dorsal pathway for visual information processing (Ungerleider & Haxby, 1994; Sheth & Young, 2016).

If this dichotomous splitting is right, it would be sensible to find a corresponding splitting in the output of visual processing, i.e., in the spatio-temporal pattern of fixations and saccades. Here, the visual assessment of tendency of the spacetime representation will prove helpful. As an example, in fig. 3, three scanpaths from the publicly available database DOVES (van der Linde, Rajashekar, Bovik, & Cormack, 2009) are shown. DOVES contains the scanpaths of 29 human observers as they viewed 101 natural im-

¹⁰Plotting a distance matrix is a technique used in different research areas and comes under different names, e.g., visual assessment of cluster tendency (VAT) (Havens, Bezdek, Keller, & Popescu, 2008; Bezdek & Hathaway, 2002), or see, e.g., Junejo, Dexter, Laptev, and Pérez (2011). In the context of dynamical systems it is called recurrence plot (Eckmann, Kamphorst, & Ruelle, 1987). Recurrence analysis is a successful tool for describing complex dynamic systems, see, e.g., Marwan, Romano, Thiel, and Kurths (2007). The reference also includes a simple statistical model for the movement of the eyes, i.e., the disrupted Brownian motion. Recurrence analysis is also known in eye movement research (Anderson, Bischof, Laidlaw, Risko, & Kingstone, 2013; Farnand, Vaidyanathan, & Pelz, 2016).

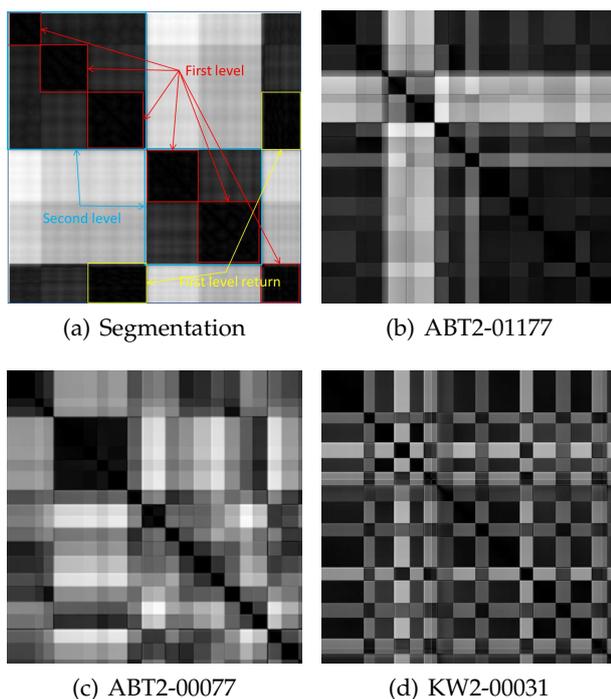


Figure 3. Hierarchy of sample clusters, first level are fixations, second level are clusters of fixations, rectangles of the first off-diagonal represent saccades

ages (van Hateren & van der Schaaf, 1998). Studying human viewing behavior while viewing pictures and images is a common subject in vision research. Since the seminal work of Buswell (1935), one often repeated general statement is that people tend to make spatially widely scattered short fixations early, transitioning to periods of spatially more confined longer fixations as viewing time increased (Babcock, Lipps, & Pelz, 2002).

This behavior is exhibited in fig. 3 (b). Here, observer CMG2 looks at stimulus img01019. Visible are three major second level blocks. The classical interpretation would be that the second block, with its more variable structure, reflects the global examining phase, while the following more homogeneous block reflects the noticing phase. The first block at the beginning represents the well known central fixation bias in scene viewing (B. W. Tatler, 2007; Bindemann, 2010).

Interestingly, the database contains also good examples for the inverse behavior, e.g., observer ABT2 looking at image img00077, see fig. 3 (c). Here the spatio-temporal pattern could be interpreted as: first the central fixation bias, second a local noticing, and only then a global scanning. This behavior is not uncommon, as Follet, Le Meur, and Baccino (2011) have noted.

These are only two examples from the database

DOVES, which contains approximately 3000 scanpaths. The visual inspection makes it possible to get a quick overview of the spatio-temporal patterns for many scanpaths and to get an intuitive understanding of prevailing pattern classes. Scanning DOVES visually shows that a significant portion of the scanpaths exhibit a spatio-temporal pattern which does not fit into the classical coarse-fine structure, e.g., subject KW2 looking at img00031 in fig. 3 (d). Of course, the examples are cursory and it is not our intention at this stage to discuss image scanning behavior. The purpose of the examples is twofold: firstly, to show that by a visual assessment of img(**D**)s, one can reach a good intuitive understanding of spatio-temporal patterns and regularities in scanpaths. The human visual system is an excellent pattern detector, a resource for investigations that should be utilized, notwithstanding the fact that a statistical examination of the data and the statistical test of hypotheses must confirm “seen” patterns. The search for simple scanpath patterns is a common task for many research questions (McClung & Kang, 2016).

Secondly, that the time course of the scanpaths is an important factor, especially when discussed in the context of top-down strategies versus bottom-up saliency. A good quantitative model should replicate the empirical observed spatio-temporal pattern classes, reflecting the order of transits between different scanning regimes and their internal substructure. The whole pattern shows a global statistics as well as substatistics in the different regimes. When modeling scanpaths, very often scanpath data are aggregated into simple feature vectors containing summary statistics as features, i.e., mean number of fixations, mean fixation duration, mean saccadic amplitude, etc. A model is considered good if it can replicate the empirical summary statistics. This neglects any time course and hierarchy in the patterns.

The next step will be to exploit the representation as a time indexed matrix of all combinatorial 2-point distances as a precise instrument of trajectory segmentation and interpretation.

Homology for spacetime coherence

At this stage, the human visual system has still been serving as pattern detector. The goal is to extract the interesting part of the information about the hierarchical spatio-temporal configuration of fixations, clusters of fixations and returns from the distance representation, and to do so on an automated basis, without any user defined parametrization, in a robust way. The question is how to express and implement this coherence algorithmically. The task will be accomplished in three steps.

Step 1: The image representation of the time indexed matrix of combinatorial 2-point distances is equivalent to a 3D surface representation, i.e., the distances are interpreted as height values over the combinatorial time \times time coordinates, see fig. 4.

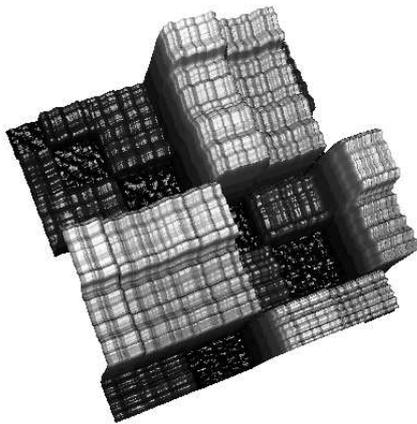
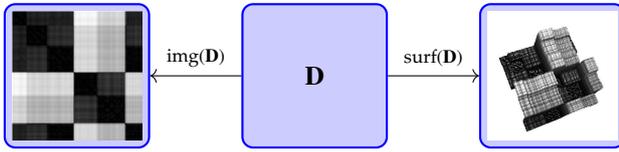


Figure 4. Surface plot of time indexed matrix of combinatorial 2-point distances

Clearly visible in the surface plot representation are rectangular columns with a small on-top variation. The small variation in blocks is considered noise. In the image view it could be regarded as a kind of texture. For a better intuitive understanding of the topological approach consider the 3D surface plot as kind of a landscape which is progressively flooded. Coherent are parts of the landscape which are below a certain sea level and form an area like a lake, without internal islands. Lying under or lying above sea level is filtering the level values according to a threshold. This is done in the next step.

Step 2: \mathbf{D} is filtered, $f_t(\mathbf{D})$, according to a threshold t , i.e., $f_t: \mathbb{R} \rightarrow \{0, 1\}$, where $\mathbf{D} \ni d_{i,j} \leq t$ then $f_t(d_{i,j}) = 1$ else $f_t(d_{i,j}) = 0$.

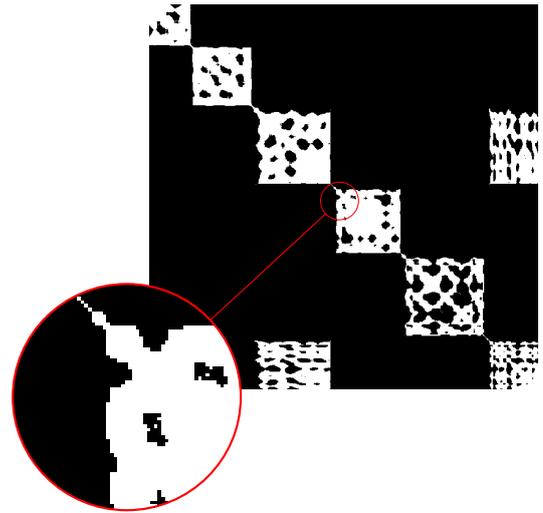
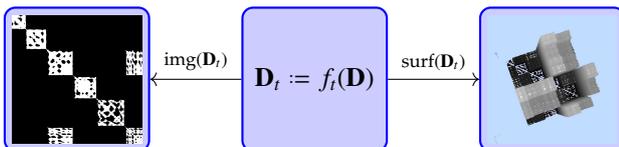
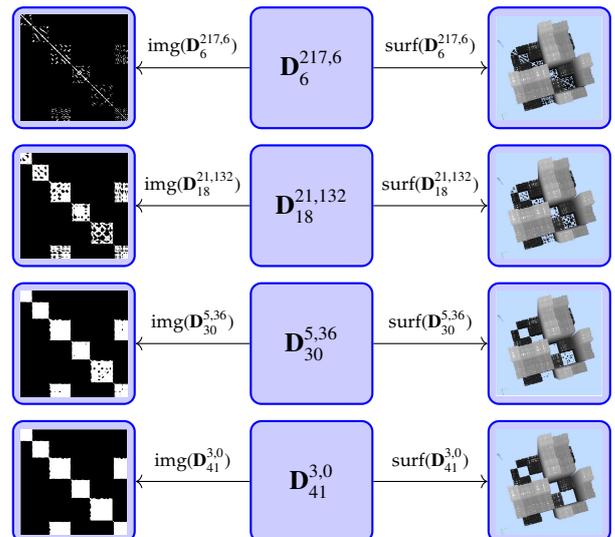


Figure 5. Filtered time indexed matrix of combinatorial 2-point distances. Magnification shows small components.

Notice the punctuated block structure in the image representation $\text{img}(f_t(\mathbf{D}))$, see fig. 5. While the overall square block structure along the diagonal and the off-diagonal rectangle block structure is still visible, the holes are representing the incoherence or noise. The incoherence is eliminated by closing the holes, i.e., raising the threshold.

Step 3: The threshold t is raised in steps from 0 until one coherent pattern along the main diagonal is formed, t_c . Coherent means that there are no holes in the connected component. The number of connected components is β_0 and the number of holes is β_1 , $\mathbf{D}_t^{\beta_0, \beta_1}$.



The coherent white part along the diagonal in the image representation is the partition of the data that we have been seeking.

It should be stated explicitly that the parameter t_c for separation is not preassigned. The definition for separation is the coherent structure/pattern of trajectory spacetime. The distance threshold is increased until coherence is reached. This is done individually for every trajectory. The pattern is global for the trajectory and does not depend on local specifics. It is important to note that a more detailed analysis within each block will separate the noise into physiological noise (tremor, drift, micro saccades, etc.) and instrument noise. In the supplementary document this approach can be interactively investigated.

All this is easy to understand for human intuition, but needs a formal mathematical theory along with an algorithm and efficient computer implementation. Generally speaking, there exist three methods to tackle the problem. The first is the obvious way, i.e., a human observer varies the “sea level”. Human evaluation especially of noisy data is common practice in eye tracking data analysis (Saez de Urabain, Johnson, & Smith, 2015). The second way is using a simple “brute force” image analysis algorithm. The third, more elegant, way is to use algebraic topology in the form of homology. Homology tells us about the connectivity and number of holes in a space, in our representation the “islands and lakes” created while flooding the space. Counting the number of connected components and the number of holes is calculating the first two Betti numbers, β_0 and β_1 , which is a fairly simple topological characteristic. The detailed description of the theory can be found in any good book on algebraic topology, e.g., Munkres (1984), Hatcher (2002), or Kaczynski, Mischaikow, and Mrozek (2010). At first sight, a formal theory might seem daunting, but the important fact is that a simple, almost trivial topological argument “no holes in trajectory spacetime” is sufficient to unambiguously determine sample clusters on different scales. The very nature of an event and a cluster of events is its “coherence” in space and time. Time comes with an order (consecutive) and space comes with a topology (vicinity, nearness).

What we have obtained is the adjacency matrix $A = [a_{i,j}]$ of graph theory for our gaze trajectory. The side length of a square around the diagonal is proportional to the duration of fixation (the time scale is fixed by the sampling rate of the gaze tracker). The rectangles in the upper and lower triangular matrix represent a return (recurrence). The length of each block contains the time information, i.e., the duration of a cluster. Separating the blocks results in

the sequence of fixations and their durations as well as the duration of intermediate gaps. Suppressing the time information in the matrix, i.e., shrinking the squares along the diagonal to one point entries, one arrives at the classical scanpath string representation of ABCDEC in the form of a matrix, see fig. 6.

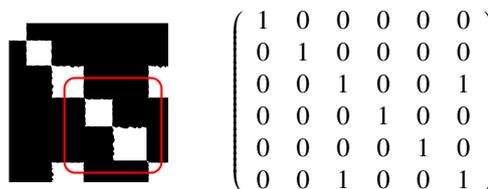


Figure 6. Matrix representation for scanpath

The off-diagonal elements are the coupling, i.e., recurrence of the fixations. The same argument for the second level squares yields the dwells, i.e., one obtains $(ABC)^1(DE)^2C^1$ (superscript numbers the dwell).

To summarize: for trajectory separation, three computational steps are needed. A distance representation for the gaze-trajectory in form of a time indexed matrix of all combinatorial 2-point distances is calculated. To separate the matrix into subparts a sliding threshold t is set, which is the sought diameter of a fixation. The threshold t is increased from 0 in steps and the number of connected parts, β_0 , and holes, β_1 , is traced. As soon as the square blocks along the diagonal form a simply connected area without holes, the minimum threshold t_c for the segmentation into fixations has been found. Further raising the threshold yields the dwells.

Abstract spacetime clustering

So far, the segmentation process for the gaze trajectory in screen space has been discussed, but the method can be made much more far-reaching. In order to do so, the meaning and interpretation of space will be generalized.

Up to now the concept of space has been the physical space and its Euclidean modeling, specifically its Euclidean metric. The crucial point is that the eyes, seen as a mechanical system, are moving in physical space, but the driving physiological and psychological processes are working in “physiological and psychological spaces”. An example of a physiological space is the color space and a much more complex space is the social space of humans when interacting, say, at a cocktail party. In this space the items or “points” are interlocutors, and the eyes are switching between these points with motivations such as signaling interest in the in-

terlocutor’s small talk, which is a gesture of politeness, and does not have the primary goal of gathering visual information. Gathering information is looking at the face to feel out the mood, etc. What counts is not the physical distance between the interlocutors, but rather some sort of social communication-distance. Relevant are the “content” of the scene and the “strategy” of the observer while interacting, which in turn is reflected in the saccade-and-fixate pattern. Physical space-distance is not a restricted resource for the eyes. The eyes can move effortlessly from one point to each other point in physical space.

As an example for the approach try for yourself the following search paradigm, see fig. 7. In the collage of colored shapes all but two colored shapes occur three times, one colored shape occurs twice and another colored shape occurs four times: which two are they? Admittedly, searching for numerosity is hard! Nevertheless, numerosity is a good example for an abstract feature, not tied to a primary sensory input. You can track and visualize your own search strategy in the supplementary interactive document.

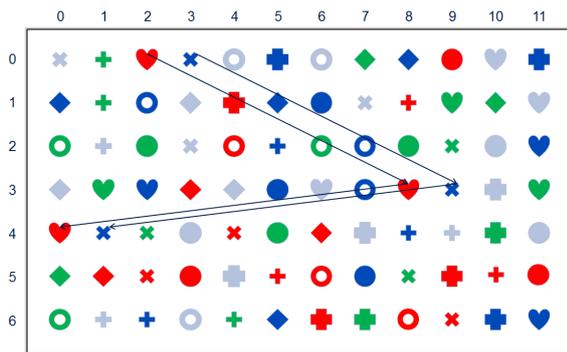


Figure 7. Search plus path

At the beginning many trajectories have fixations on a color. This derives from the fact that humans can identify color-blobs very easily in their view field. Thus, the first “search channel” is very often color.¹¹ The second channel is an easily detectable “geometry”. While the distinct color blobs are far apart in terms of geometric Euclidean distance they are near in color-space, i.e., the red disk (0,9) is near, actually identical, in color to the red disks (5,3) and (5,11). The same holds true for the “geometry channel”, e.g., the motives with a circular boundary. It is likely that most subjects will start out with a random search strategy, which after a while will be abandoned in favor of a systematic, row-by-row, search strategy.

The qualitative approach to the geometric stimuli analysis is taken in “Gestaltpsychology”. A more recent and formal approach to it is taken in structural

information theory and algorithmic information theory, which can be made quantitative. Using specialized metrics differentiates the channels in the search strategy in a metric way and helps to classify viewers. It is helpful to change the terminology and to say that the eyes are moving in “feature space”. This space has different dimensions like color, shape, etc., which form subspaces. The feature space is a topological space. For ease of use it could be modeled as a metric space and the path is encoded in feature distance. Of course, the metric has to be adapted for special purposes. A simple example is the distance in color-space. Simple is certainly relative, taking into account the long way from first color theories of the 19th century into the elaborated color spaces like the HUE space, used in printing and computer imaging. This development has by no means come to an end. A (much) more complex example is the distance in social interaction.

Nevertheless, the starting point is always the basic notion of a metrizable “neighborhood or nearness” relation in the form of a metric. The metric is the crucial starting point to emphasize different aspects in the trajectory. Let us start with the metric on a space X . The general mathematical notion of a metric is a function

$$d(x, y) : X \times X \rightarrow \mathbb{R}$$

satisfying for all $x, y, z \in X$ the conditions

Positiveness: $d(x, y) \geq 0$ with equality only for $x = y$

Symmetry: $d(x, y) = d(y, x)$

Triangle inequality: $d(x, y) \leq d(x, z) + d(z, y)$

This definition is only the bare skeleton of a metric. By itself it does not preassign any structure in the data, as is shown in the example:

$$d(P_i, P_j) = 0 \text{ for } i = j \text{ else } 1$$

A more complex metric gives a much richer structure, emphasizing interesting aspects in the data. In RGB color space the distance between two colors $C_1(R, G, B)$ and $C_2(R, G, B)$ simply is:

$$d_{RGB}(C_1, C_2) = \sqrt{(R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2}$$

A different example is reading. Here it would be appropriate to work within text space. For the understanding of reading patterns, not only the physical spacing of characters, but also the semantic distance is important. The semantic distance measures the difficulty of

¹¹Anatomically a separate pathways for color can be distinguished (Schwartz, 2010).

understanding words in a reading context. In the flow of reading, words can be physically close together, but if a word does not fit into the context or is not known to the reader, the reader will have difficulties in processing the word and a regression is most likely. Understanding a text requires coherence of word semantics as well as with the narrative in which they occur. The reader is traveling in general feature spaces and coherence is maintained or broken.

Along these lines more complex spaces can be constructed and analyzed. Clustering the data in feature space reveals directly the process related time ordering without intermediate separation of data into fixations, saccades, and then assigning areas of interest. The process pattern works directly on the items of interest. To cite Stark and Ellis (1981)

Sensory elements are semantic subfeatures of scenes or pictures being observed and motor elements are saccades that represent the syntactical structural or topological organization of the scene.

The ITop algorithm is essentially meant for stimulus-space based analyses. The idea of directly connecting stimuli information and eye tracking data is also proposed in (Andersson et al., 2016).

Results for fixation identification

To show the algorithm's potential for level one eye-tracking data segmentation, a basic comparison with a state-of-the-art algorithm is given. An in-depth evaluation together with a MATLAB® reference implementation will be provided in a follow-up article. Current research has raised the awareness that algorithms commonly in use, especially when used "out of the box", markedly differ in their results and an overall standard is lacking (Andersson et al., 2016). This situation escalates with each new algorithm proposed. The topological approach introduced herein is no exception. To make results comparable as much as possible a common reference set together with computed results, e.g., number and duration of events, event detected at samples, would be preferable. In a recent article, (Hessels et al., 2016) introduced a new algorithm, identification by two-means clustering (I2MC), together with an open source reference implementation as well as ten datasets to show the performance of their approach. The I2MC algorithm is evaluated against seven state-of-the-art event detection algorithms and is reported to be the most robust to high noise and data loss levels, which makes it suitable for eye-tracking research with infants, school children, and

certain patient groups. To ensure performance and comparability the identification by topological arguments (ITop) is checked against I2MC. The data are taken from www.github.com/royhessels/I2MC. The datasets comprise two participants, each participant having five trials, resulting in ten datasets overall. Both eyes are tracked. I2MC makes use of the data from both eyes for fixation detection, ITop classifies solely on the basis of the left eye data series. I2MC uses an interpolation algorithm for gap-filling. ITop works without gap filling. Fig. 8 shows the classification results for the ten datasets under the ITop and I2MC algorithm.

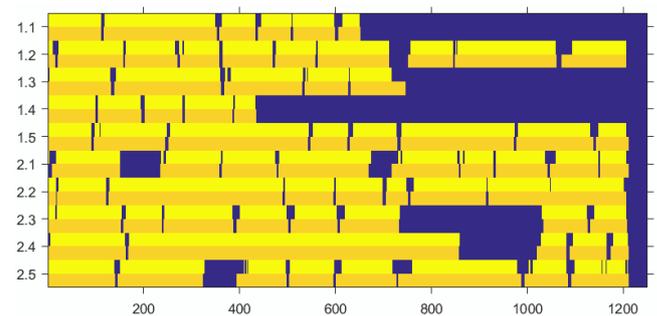


Figure 8. Performance of ITop and I2MC on ten datasets. The y-axis is in participant.trial, the x-axis is in samples. ITop fixation periods are in yellow and I2MC fixation periods are in orange. Dark blue is the gap between detected fixations or periods of data loss.

At some positions the ITop signal is splitted into two peaks, e.g., 1.3 (at samples 360–382 and 533–542) and 2.5 (at samples 1155–1165). This is no error, it is a finer view of the data. This is discussed in the following examples. The two approaches are in good agreement. Whenever I2MC detects a fixation ITop also does. ITop detects two additional fixations, one for 2.2 (at samples 1048–1049) and one for 2.3 (at samples 17–19). A closer look at the scatter plot as well as the position plot reveals two very close fixations, see (fig. 9, fig. 10) and (fig. 11, fig. 12).

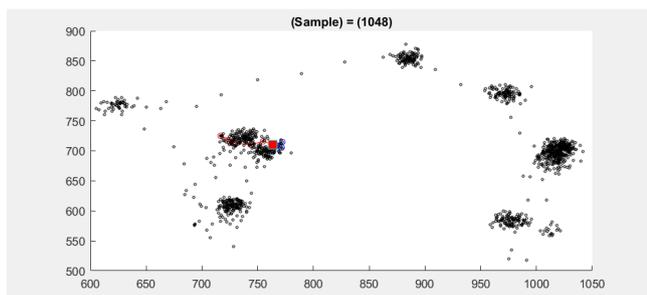


Figure 9. Scatter plot for dataset 2.2 at sample 1048 (red square at sample 1048) shows two clusters very close to each other.

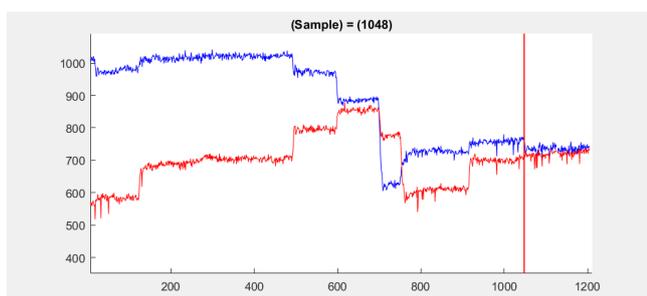


Figure 10. Position plot for dataset 2.2 at sample 1048 (red line at sample 1048) shows a small jump in the mean. The small jump is detected in spite of significant noise.

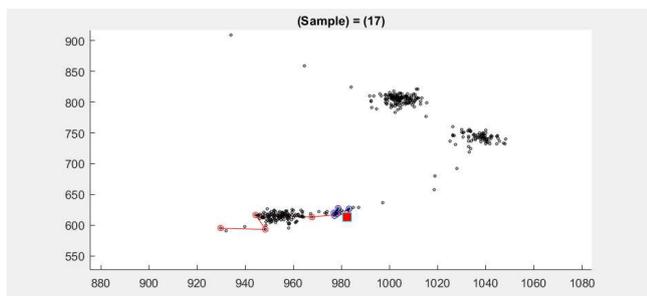


Figure 11. Scatter plot for dataset 2.3 at samples 17–19 (red square at sample 18) shows two clusters.

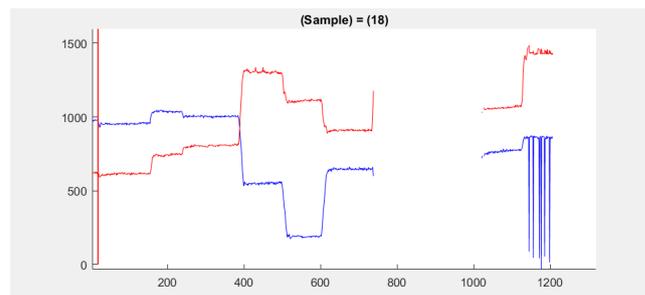


Figure 12. Position plot for dataset 2.3 at samples 17–19 (red line at sample 18) shows a small jump in the mean.

Although no data interpolation is done, ITop can identify a shift in the direct neighborhood of data loss. This is shown for 2.1 at samples 242–246, see fig. 13.

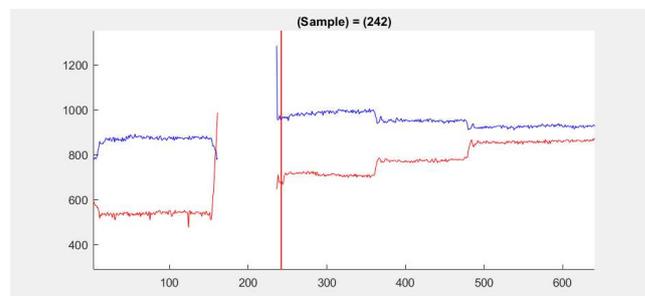


Figure 13. Position plot for dataset 2.1 at samples 242–246 (red line at sample 242) shows a small jump in the mean after a period of data loss.

At some positions the gap between fixations is split, e.g., for 1.3 at samples 360–382. This is a finer view of the data. As discussed, a saccade very often shows a complex stopping signal (Hooge, Nyström, Cornelissen, & Holmqvist, 2015), post saccadic oscillations are a prominent example (Nyström & Holmqvist, 2010). The term complex is meant in contrast to abrupt stopping. It does not necessarily mean a post-saccadic oscillation (PSO). A PSO is only an example for a named event with a more complicated “braking” pattern. This is reflected in the splitting of the signal. The position plot for 1.3 at samples 360–382 shows such a complex behavior, see fig. 14.

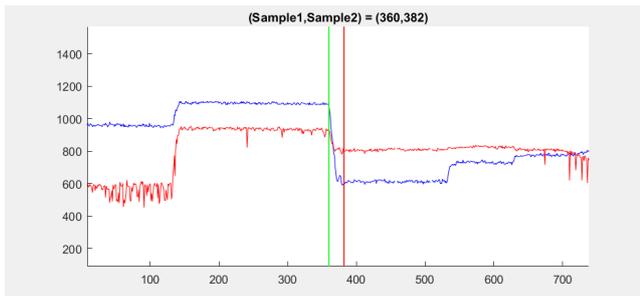


Figure 14. Position plot for dataset 1.3 between sample 360 (green line) and sample 382 (red line) showing a complex transit between two fixations.

The splitting according to braking can be much finer but is still detected by ITop. An example is 1.3 at samples 533–543. Here, a very small shift in the mean of the y-position signal occurs shortly after stopping, showing the high sensitivity of ITop, see fig. 15.

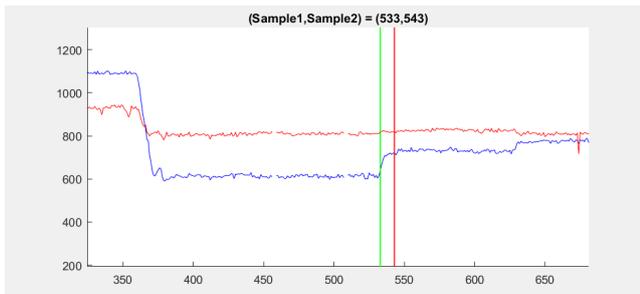


Figure 15. Position plot for dataset 1.3 between sample 533 (green line) and sample 543 (red line) showing a small jump in the mean of the y-position after stopping. The jump occurs at the red line.

It must further be noted that the saccades according to ITop are longer (spatially wider) than under I2MC. As an example, dataset 2.3 at samples 499–515 is shown in detail. I2MC detects a gap between two fixations at samples 502–507, see fig. 16.

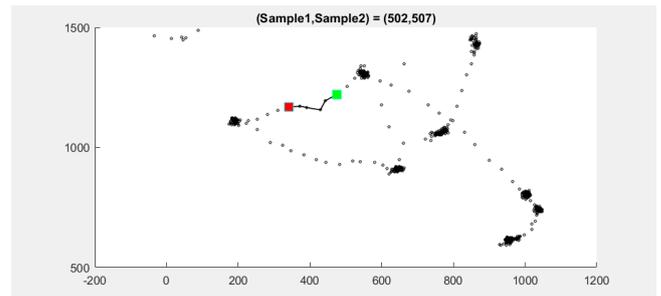


Figure 16. Scatter plot for dataset 2.3 between sample 502 (green square) and sample 507 (red square).

ITop detects the gap at the same location at samples 499–515 and is therefore approximately twice as long, see fig. 17.

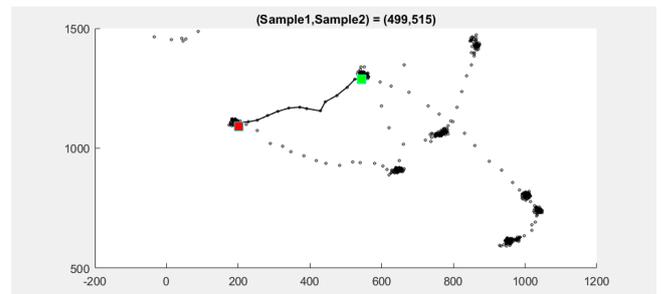


Figure 17. Scatter plot for dataset 2.3 between sample 499 (green square) and sample 515 (red square).

The position plot shows a jag in the y-signal, which could potentially mislead an algorithm, see fig. 18.

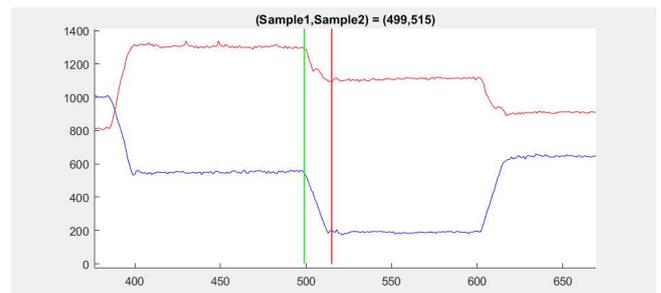


Figure 18. Position plot for dataset 2.3 between sample 499 (green line) and sample 515 (red line). A jag occurs at sample 504, potentially misleading algorithms.

ITop also indicates other changes in the data series, like stationarity, e.g., the double peaked signal for dataset 2.5 at samples 1155–1165 indicates the onset of a drift in a fixation, see fig. 19.

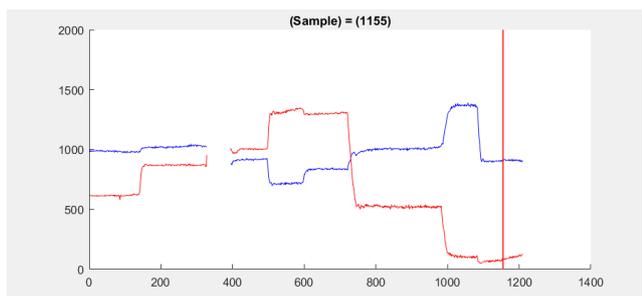


Figure 19. Position plot for dataset 2.5 shows a drift beginning at sample 1155 (red line).

Notwithstanding that I2MC and ITop are in good overall agreement they also show differences on a finer scale. If one takes into consideration the broad number of algorithms and different approaches for event detection it must be clear that the overall results can be markedly different. This can only be mitigated by defining events in an unambiguous and definite way and comparing algorithms on the basis of standard data on a sample by sample level.

Discussion

A general overview of the algorithms currently in use for event detection in eye-tracking data is given, showing that there is no standard for event detection, even in the case of the most basic events such as fixations and saccades.

A topological approach to event detection in raw eye-tracking data is introduced, ITop. The detection is based on the topological abstraction of coherence in space and time of the sample points. The idea of trajectory spacetime coherence is given a precise meaning in topological terms, i.e., “no holes in trajectory spacetime”, a strikingly simple topological argument for the separation of the sample data. The topological argument is a kind of common rationale for most of the algorithms currently in use. The basis for the topological approach is the representation of raw eye-tracking data in the form of a time indexed matrix of combinatorial 2-point distances. This representation makes the coherence of sample data in space and time easily accessible. The time ordered 2-point combinatorial distances representation makes the gaze trajectory independent of Euclidean motions, which is a desired property when comparing scanpaths, since distances are the invariants of Euclidean geometry.

For visualization, the matrix is displayed as a grayscale image to show the spatio-temporal ordering and coherence of the gaze-points in display space.

For the human visual system the interesting parts are easy to detect, e.g., fixations, dwells, etc. The visual assessment of spatio-temporal coherence is discussed and exemplified in the context of coarse-fine (global-local) scanpath characteristics. It is argued that the visual assessment of the trajectory spacetime is helpful to identify general patterns in viewing behavior and to develop an intuitive understanding thereof.

To separate fixations and higher level clusters of fixations out of eye-tracking data, the common argument of spatio-temporal coherence, implicitly used in existing algorithms, is converted into an explicit topological argument, i.e., “no holes in trajectory spacetime”. The method encompasses the well known criteria which are partially expressed as thresholds for velocity, acceleration, amplitude, duration, etc. Tracking the number of connected parts and holes while varying the scale allows the partitioning of the distances matrix into the classical scanpath oculomotor events, i.e., segments of fixations and saccades. The segments are identified by their spatio-temporal coherence by means of simple homology, which is a classical tool of algebraic topology. For processing the data no preprocessing is needed, i.e., gap-filling, filtering, and smoothing, preserving the data “as is”. This approach makes it possible to identify the single events without any predefined parameters. A postprocessing of the found events, like merging of nearby fixations or the removal of physiologically implausible short fixations and saccades is not needed.

The topological segmentation is introduced in the familiar setting of Euclidean space and its well known metric. The advantage of this approach is that it can be easily expanded to general spaces like color spaces, shape spaces, etc., allowing the analysis of complex patterns in higher human activities. The ITop algorithm is essentially meant for stimuli-space based analysis.

In order to facilitate the intuitive understanding the article is accompanied by a supplementary interactive document.

ITop is considered as a fourth approach to eye-tracking data in addition to the well known threshold based approaches and the newer probabilistic and machine learning methods. An expanded comparison, analysis, and classification of the ITop detection patterns together with an open source MATLAB[®] reference implementation will be provided in a further work.

Acknowledgement

We thank the anonymous reviewers who provided helpful comments on earlier drafts of the manuscript

and whose comments/suggestions helped to improve and clarify this manuscript. The provision of important references and preprints is also greatly appreciated.

References

- Anantrasirichai, N., Gilchrist, I. D., & Bull, D. R. (2016). Fixation identification for low-sample-rate mobile eye trackers. In *2016 IEEE International Conference on Image Processing (ICIP)* (p. 3126-3130). doi: 10.1109/ICIP.2016.7532935
- Anderson, N. C., Bischof, W. F., Laidlaw, K. E. W., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior Research Methods*, 45(3), 842–856.
- Andersson, R., Larsson, L., Holmqvist, K., Stridh, M., & Nyström, M. (2016). One algorithm to rule them all? an evaluation and discussion of ten eye movement event-detection algorithms. *Behavior Research Methods*, 1–22. Retrieved from <http://dx.doi.org/10.3758/s13428-016-0738-9> doi: 10.3758/s13428-016-0738-9
- Andersson, R., Nyström, M., & Holmqvist, K. (2010). Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more. *Journal of Eye Movement Research*, 3(3), 1–12.
- Artal, P. (2014, Sep). Optics of the eye and its impact in vision: a tutorial. *Adv. Opt. Photon.*, 6(3), 340-367.
- Arzi, M., & Magnin, M. (1989). A fuzzy set theoretical approach to automatic analysis of nystagmic eye movements. *IEEE Transactions on Biomedical Engineering*, 36(9), 954-963.
- ASL. (2007). Eye tracker system manual asl eyetrac 6 eyenal analysis software [Computer software manual].
- Babcock, J. S., Lipps, M., & Pelz, J. B. (2002). How people look at pictures before, during, and after scene capture: Buswell revisited. In *Proc. SPIE 4662, human vision and electronic imaging VII* (Vol. 4662, p. 34-47). Retrieved from <http://dx.doi.org/10.1117/12.469552> doi: 10.1117/12.469552
- Bahill, A. T. (1983). Predicting final eye position halfway through a saccade. *IEEE Transactions on Biomedical Engineering*, 30(12), 781–786.
- Bahill, A. T., Brockenbrough, A., & Troost, B. T. (1981). Variability and development of a normative data base for saccadic eye movements. *Invest. Ophthalmol. Vis. Sci.*, 21(1), 116–125.
- Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences*, 24, 191–204.
- Behrens, F., MacKeben, M., & Schröder-Preikschat, W. (2010). An improved algorithm for automatic detection of saccades in eye movement data and for calculating saccade parameters. *Behavior Research Methods*, 3, 701–708.
- Behrens, F., & Weiss, L.-R. (1992). An algorithm separating saccadic from nonsaccadic eye movements automatically by use of the acceleration signal. *Vision Research*, 32(5), 889 - 893.
- Bennett, A., & Rabbetts, R. B. (2007). *Clinical visual optics* (fourth ed.). Butterworth Heinemann Elsevier.
- Berg, D. J., Boehnke, S. E., Marino, R. A., Munoz, D. P., & Itti, L. (2009). Free viewing of dynamic stimuli by humans and monkeys. *Journal of Vision*, 9(5), 1–15.
- Bezdek, J., & Hathaway, R. (2002). Vat: a tool for visual assessment of (cluster) tendency. In *Neural networks, 2002. IJCNN '02. proceedings of the 2002 international joint conference on* (pp. 2225–2230).
- Bindemann, M. (2010). Scene and screen center bias early eye movements in scene viewing. *Vision Research*, 50(23), 2577 - 2587. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0042698910004025> (Vision Research Reviews) doi: <http://dx.doi.org/10.1016/j.visres.2010.08.016>
- Blackmon, T. T., Ho, Y. F., Chernyak, D. A., Azzariti, M., & Stark, L. W. (1999). Dynamic scanpaths: Eye movement analysis methods. In *IS&T/SPIE conference on human vision and electronic imaging IV SPIE vol. 3644*.
- Blignaut, P. (2009). Fixation identification: the optimum threshold for a dispersion algorithm. *Attention, Perception, & Psychophysics*, 71(4), 881–895.
- Bollen, E., Bax, J., vanDijk, J. G., Koning, M., Bos, J. E., Kramer, C. G. S., & van der Velde, E. A. (1993). Variability of the main sequence. *Investigative Ophthalmology & Visual Science*, 34(13), 3700–3704.
- Borji, A., & Itti, L. (2014). Defending yarbus: Eye movements reveal observers' task. *Journal of Vision*, 14(3), 29.
- Brasel, S. A., & Gips, J. (2008). Points of view: Where do we look when we watch tv? *Perception*, 37, 1890–1894.
- Buhmann, C., Maintz, L., Hierling, J., Vettorazzi, E., Moll, C. K., Engel, A. K., ... Zangemeister, W. H. (2014). Effect of subthalamic nucleus deep brain stimulation on driving in parkinson disease. *Neurology*, 82(1), 32–40. Retrieved from <http://www.neurology.org/content/early/2013/12/18/01.wnl.0000438223.17976.fb.abstract> doi:

- 10.1212/01.wnl.0000438223.17976.fb
- Buswell, G. T. (1935). *How people look at pictures: A study of the psychology of perception in art*. The University of Chicago Press.
- Camilli, M., Nacchia, R., Terenzi, M., & Nocera, F. D. (2008). Astef: A simple tool for examining fixations. *Behavior Research Methods*, 40(2), 373–382.
- Carmi, R., & Itti, L. (2006). The role of memory in guiding attention during natural vision. *Journal of Vision*, 6, 898–914.
- Carpenter, R. H. S. (1988). *Movements of the eyes* (2nd ed.). Pion.
- Crabb, D. P., Smith, N. D., Rauscher, F. G., Chisholm, C. M., Barbur, J. L., Edgar, D. F., & Garway-Heath, D. F. (2010, 03). Exploring eye movements in patients with glaucoma when viewing a driving scene. *PLoS ONE*, 5(3), e9710.
- Cuong, N. V., Dinh, V., & Ho, L. S. T. (2012). Mel-frequency cepstral coefficients for eye movement identification. In *2012 IEEE 24th International Conference on Tools with Artificial Intelligence* (p. 253-260). doi: 10.1109/ICTAI.2012.42
- Czabanski, R., Pander, T., & Przybyla, T. (2014). Fuzzy approach to saccades detection in optokinetic nystagmus. In D. A. Gruca, T. CzachÅrski, & S. Kozielski (Eds.), *Man-machine interactions 3* (Vol. 242, p. 231-238). Springer International Publishing.
- Daye, P. M., & Optican, L. M. (2014). Saccade detection using a particle filter. *Journal of Neuroscience Methods*, 235, 157 - 168.
- de Bruin, J. A., Malan, K. M., & Eloff, J. H. P. (2013). Saccade deviation indicators for automated eye tracking analysis. In *Proceedings of the 2013 conference on eye tracking south africa* (p. 47-54). New York, NY, USA: ACM.
- Dorr, M., Jarodzka, H., & Barth, E. (2010). Space-variant spatio-temporal filtering of video for gaze visualization and perceptual learning. In *Etra '10: Proceedings of the 2010 symposium on eye-tracking research & applications* (pp. 307–314). New York, NY, USA: ACM. doi: 10.1145/1743666.1743737
- Dorr, M., Vig, E., & Barth, E. (2012). Eye movement prediction and variability on natural video data sets. *Visual Cognition*, 1–20.
- Duchowski, A. T. (1998, March). 3d wavelet analysis of eye movements. In H. H. Szu (Ed.), *Proc. spie 3391, wavelet applications v*, 435.
- Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, 34(4), 455–470.
- Duchowski, A. T. (2007). *Eye tracking methodology* (2nd ed.). Springer.
- Eckmann, J.-P., Kamphorst, S. O., & Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhysics Letters*, 4(9), 973–977.
- Elbaum, T., Wagner, M., & Botzer, A. (2017). Cyclopean vs. dominant eye in gaze-interface-tracking. *Journal of Eye Movement Research*, 10. Retrieved from <https://bop.unibe.ch/index.php/JEMR/article/view/2961>
- Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, 43, 1035–1045.
- Engbert, R., & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences of the United States*, 103(18), 7192–7197.
- Engbert, R., Mergenthaler, K., Sinn, P., & Pikovsky, A. (2011). An integrated model of fixational eye movements and microsaccades. *Proceedings of the National Academy of Sciences of the United States of America*, 108(3).
- Farnand, S., Vaidyanathan, P., & Pelz, J. (2016). Recurrence metrics for assessing eye movements in perceptual experiments. *Journal of Eye Movement Research*, 9(4). doi: 10.16910/jemr.9.4.1
- Fischer, B., Biscaldi, M., & Otto, P. (1993). Saccadic eye movements of dyslexic adult subjects. *Neuropsychologia*, 31(9), 887-906.
- Follet, B., Le Meur, O., & Baccino, T. (2011). New insights into ambient and focal visual fixations using an automatic classification algorithm. *i-Perception*, 2(6). Retrieved from <https://doi.org/10.1068/i0414> doi: 10.1068/i0414
- Frey, H.-P., Honey, C., & König, P. (2008). What's color got to do with it? the influence of color on visual attention in different categories. *Journal of Vision*, 8(14).
- Gilchrist, I. D. (2011). The oxford handbook of eye movements. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), (chap. Saccades). Oxford University Press.
- Gitelman, D. R. (2002). Ilab: A program for postexperimental eye movement analysis. *Behavior Research Methods, Instruments, & Computers*, 34(4), 605 - 612.
- Godwin, H. J., Reichle, E. D., & Menneer, T. (2014). Coarse-to-fine eye movement behavior during visual search. *Psychonomic Bulletin & Review*, 21(5), 1244–1249. Retrieved from <http://dx.doi.org/10.3758/s13423-014-0613-6> doi: 10.3758/s13423-014-0613-6
- Goldberg, J. H., & Schryver, J. C. (1995). Eye-gaze-contingent control of the computer interface: Methodology and example for zoom detec-

- tion. *Behavior Research Methods*, 27(3), 338–350.
- Goldberg, J. H., & Wichansky, A. M. (2003). Eye tracking in usability evaluation: A practitioner's guide. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 493–516). North-Holland.
- Goldstein, R. B., Woods, R. L., & Peli, E. (2007). Where people look when watching movies: Do all viewers look at the same place? *Comput Biol Med.*, 37(7), 957–964.
- Green, P. (2002). Where do drivers look while driving (and for how long)? In R. Dewar & P. Olson (Eds.), *Human factors in traffic safety* (2nd ed., p. 57-82). Lawyers & Judges.
- Greene, M. R., Liu, T., & Wolfe, J. M. (2012). Reconsidering yarbus: A failure to predict observers' task from eye movement patterns. *Vision Research*, 62, 1 - 8. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0042698912000922> doi: <http://dx.doi.org/10.1016/j.visres.2012.03.019>
- Groner, R., & Groner, M. (1982). Towards a hypothetico-deductive theory of cognitive activity. In R. Groner & P. Fraisse (Eds.), *Cognition and eye movements* (p. 100-121). Retrieved from https://www.researchgate.net/publication/312424385_Groner_R_Groner_M_1982_Towards_a_hypothetico-deductive_theory_of_cognitive_activity_In_R_Groner_P_Fraisse_Eds_Cognition_and_eye_movements_Amsterdam_North_Holland
- Groner, R., & Groner, M. T. (1989). Attention and eye movement control: An overview. *European archives of psychiatry and neurological sciences*, 239(1), 9–16. Retrieved from <http://dx.doi.org/10.1007/BF01739737> doi: 10.1007/BF01739737
- Groner, R., Walder, F., & Groner, M. (1984). Looking at faces: Local and global aspects of scanpaths. In A. G. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of eye movement research selected/edited proceedings of the second european conference on eye movements* (Vol. 22, p. 523 - 533). North-Holland. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0166411508618749> doi: [http://dx.doi.org/10.1016/S0166-4115\(08\)61874-9](http://dx.doi.org/10.1016/S0166-4115(08)61874-9)
- Gustafsson, F. (2000). *Adaptive filtering and change detection*. Wiley.
- Haji-Abolhassani, A., & Clark, J. J. (2014). An inverse yarbus process: Predicting observers' task from eye movement patterns. *Vision Research*, 103, 127 - 142. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0042698914002004> doi: <http://dx.doi.org/10.1016/j.visres.2014.08.014>
- Hatcher, A. (2002). *Algebraic topology*. Cambridge University Press.
- Havens, T., Bezdek, J., Keller, J., & Popescu, M. (2008, dec.). Dunn's cluster validity index as a contrast measure of vat images. In *Pattern recognition, 2008. icpr 2008. 19th international conference on* (pp. 1–4).
- Helo, A., Rämä, P., Pannash, S., & Meary, D. (2016). Eye movement patterns and visual attention during scene viewing in 3- to 12-month-olds. *Visual Neuroscience*, 33. Retrieved from <https://www.cambridge.org/core/article/div-class-title-eye-movement-patterns-and-visual-attention-during-scene-viewing-in-3-to-12-month-olds-div/625327B87FEED9F7771B929A94926C53> doi: 10.1017/S0952523816000110
- Hessels, R. S., Niehorster, D. C., Kemner, C., & Hooge, I. T. C. (2016). Noise-robust fixation detection in eye movement data: Identification by two-means clustering (i2mc). *Behavior Research Methods*, 1–22. Retrieved from <http://dx.doi.org/10.3758/s13428-016-0822-1> doi: 10.3758/s13428-016-0822-1
- Holmqvist, K., Nyström, M., Anderson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking*. Oxford University Press.
- Hooge, I., Nyström, M., Cornelissen, T., & Holmqvist, K. (2015). The art of braking: Post saccadic oscillations in the eye tracker signal decrease with increasing saccade size. *Vision Research*, 112, 55 - 67. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0042698915001170> doi: <http://dx.doi.org/10.1016/j.visres.2015.03.015>
- Hoppe, S., & Bulling, A. (2016). End-to-end eye movement detection using convolutional neural networks. *arXiv e-prints*. Retrieved from <http://arxiv.org/abs/1609.02452>
- Horn, A. K., & Adamczyk, C. (2012). Chapter 9 - reticular formation: Eye movements, gaze and blinks. In J. K. M. Paxinos (Ed.), *The human nervous system (third edition)* (Third Edition ed., p. 328 - 366). San Diego: Academic Press.
- Inchingolo, P., & Spanio, M. (1985). On the identification and analysis of saccadic eye movements—a quantitative study of the processing procedures. *Biomedical Engineering, IEEE Transactions on*, 32(9), 683–695.

- Inhoff, A. W., Seymoura, B. A., Schad, D., & Greenberg, S. (2010). The size and direction of saccadic curvatures during reading. *Vision Research*, 50(12), 1117–1130.
- Itti, L. (2005). Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition*, 12, 1093–1123.
- Jarodzka, H., Holmqvist, K., & Nyström, M. (2010). A vector-based, multidimensional scanpath similarity measure. In *Etra '10: Proceedings of the 2010 symposium on eye-tracking research & applications* (pp. 211–218). New York, NY, USA: ACM. doi: <http://doi.acm.org/10.1145/1743666.1743718>
- Junejo, I. N., Dexter, E., Laptev, I., & Pérez, P. (2011). View-independent action recognition from temporal self-similarities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(1), 172–185.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, 87(4), 329–354.
- Jüttner, M., & Wolf, W. (1992). Occurrence of human express saccades depends on stimulus uncertainty and stimulus sequence. *Experimental Brain Research*, 89(3), 678–681.
- Kaczynski, T., Mischaikow, K., & Mrozek, M. (2010). *Computational homology* (S. S. Antmann, J. E. Marsden, & S. L., Eds.) (No. 157). Springer.
- Karn, K. S. (2000, November). “saccade pickers” vs. “fixation pickers”: The effect of eye tracking instrumentation on research. In *Proceedings eye tracking research & applications symposium 2000* (pp. 87–88).
- Karsh, R., & Breitenbach, F. W. (1983). Looking at looking: The amorphous fixation measure. In R. Groner (Ed.), *Eye movements and psychological functions: International views* (pp. 53–64). Hillsdale: Lawrence Erlbaum.
- Kasneci, E., Kasneci, G., Kübler, T. C., & Rosenstiel, W. (2014). The applicability of probabilistic methods to the online recognition of fixations and saccades in dynamic scenes. In *Proceedings of the symposium on eye tracking research and applications* (pp. 323–326). New York, NY, USA: ACM.
- Kliegl, R., & Olson, R. K. (1981). Reduction and calibration of eye monitor data. *Behavior Research Methods & Instrumentation*, 13(2), 107–111.
- Koh, D. H., Gowda, S. M., & Komogortsev, O. V. (2010). Real time eye movement identification protocol. In *Extended abstracts of the 28th international conference on human factors in computing systems* (pp. 3499–3504).
- Komogortsev, O. V., Gobert, D. V., Jayarathna, S., Koh, D. H., & Gowda, S. M. (2010). Standardization of automated analyses of oculomotor fixation and saccadic behaviors. *IEEE Transactions on Biomedical Engineering*, 57(11), 2635–2645.
- Komogortsev, O. V., Jayarathna, S., Koh, D. H., & Gowda, S. M. (2009). *Qualitative and quantitative scoring and evaluation of the eye movement classification algorithms* (Tech. Rep.). Texas State University San Marcos Department of Computer Science.
- Komogortsev, O. V., & Karpov, A. (2012/13). Automated classification and scoring of smooth pursuit eye movement in presence of fixation and saccades. *Behavior Research Methods*, 45(1), 203–215.
- Komogortsev, O. V., & Khan, J. I. (2007). Kalman filtering in the design of eye-gaze-guided computer interfaces. In *Human-computer interaction. hci intelligent multimodal interaction environments: 12th international conference, hci international 2007, beijing*.
- Krassanakis, V., Filippakopoulou, V., & Nakos, B. (2014). Eyemv toolbox: An eye movement post-analysis tool based on a two-step spatial dispersion threshold for fixation identification. *Journal of Eye Movement Research*, 7(1), 1–10.
- Krauzlis, R. J., & Miles, F. A. (1996). Release of fixation for pursuit and saccades in humans: Evidence for shared inputs acting on different neural substrates. *Journal of Neurophysiology*, 76(5), 2822–2833.
- Kumar, M., Klingner, J., Puranik, R., Winograd, T., & Paepcke, A. (2008). Improving the accuracy of gaze input for interaction. In *Etra '08: Proceedings of the 2008 symposium on eye tracking research & applications* (pp. 65–68). ACM.
- Land, M. F. (2011). Oculomotor behavior in vertebrates and invertebrates. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The oxford handbook of eye movements* (chap. 1). Oxford University Press.
- Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Phil. Trans. R. Soc. Lond. B*, 352(1358), 1231–1239.
- Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature*, 369, 742–744.
- Land, M. F., & Tatler, B. W. (2009). *Looking and acting*. Oxford University Press.
- Lappi, O., & Lehtonen, E. (2013). Eye-movements in real curve driving: pursuit-like optokinesis in vehicle frame of reference, stability in an allocentric reference coordinate system. *Journal of Eye Movement Research*, 6(1), 1–13.
- Larsson, P. (2002). *Automatic visual behavior analysis*

- (Unpublished master's thesis). Control and Communication Department of electrical engineering Linköping University, Sweden.
- Leigh, R. J., & Kennard, C. (2004, Mar). Using saccades as a research tool in the clinical neurosciences. *Brain*, 127(Pt 3), 460–477. Retrieved from <http://dx.doi.org/10.1093/brain/awh035> doi: 10.1093/brain/awh035
- Leigh, R. J., & Zee, D. S. (2006). *The neurology of eye movements* (4th ed.). Oxford.
- Liston, D. B., Krukowski, A. E., & Stone, L. S. (2012). Saccade detection during smooth tracking. *Displays*.
- Liversedge, S. P., Gilchrist, I. D., & Everling, S. (Eds.). (2011). *The oxford handbook of eye movements*. Oxford University Press.
- Longbotham, H. G., Engelken, E. J., Rea, J., Shelton, D., R., R., A., C., & Harris, J. (1994). Nonlinear approaches for separation of slow and fast phase nystagmus signals. *Biomedical Sciences Instrumentation*, 30, 99–104.
- Manor, B. R., & Gordon, E. (2003). Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. *Journal of Neuroscience Methods*, 128(1-2), 85-93.
- Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Hubel, D. H. (2009). Microsaccades: a neurophysiological analysis. *Trends in Neurosciences*, 32(9), 463-475.
- Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5–6), 237–329.
- Mason, R. L. (1976). Digital computer estimation of eye fixations. *Behavior Research Methods & Instrumentation*, 8(2), 185-188.
- Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin*, 81(12), 899-917.
- Matsuoka, K., & Harato, H. (1983). Detection of rapid phases of eye movements using third order derivatives. *Japanese J. Ergonomics*, 19, 147–153.
- McClung, S. N., & Kang, Z. (2016). Characterization of visual scanning patterns in air traffic control. *Computational Intelligence and Neuroscience*. Retrieved from <https://www.hindawi.com/journals/cin/2016/8343842/> doi: 10.1155/2016/8343842
- Menz, C., & Groner, R. (1985). The effects of stimulus characteristics, task requirements and individual differences on scanning patterns. In R. Groner, G. W. McConkie, & C. Menz (Eds.), *Eye movements and human information processing*. Amsterdam: North Holland.
- Mosquera, S., Verma, S., & McAlinden, C. (2015). Centration axis in refractive surgery. *Eye and Vision*, 2(1), 4.
- Mould, M. S., Foster, D. H., Amano, K., & Oakley, J. P. (2012). A simple nonparametric method for classifying eye fixations. *Vision Research*, 57(15), 18–25.
- Munkres, J. R. (1984). *Elements of algebraic topology*. Perseus Publishing.
- Munn, S. M., Stefano, L., & Pelz, J. B. (2008). Fixation-identification in dynamic scenes: comparing an automated algorithm to manual coding. In *Apgv '08: Proceedings of the 5th symposium on applied perception in graphics and visualization* (pp. 33–42). New York, NY, USA: ACM.
- Munoz, D. P., Armstrong, I., & Coe, B. (2007). Using eye movements to probe development and dysfunction. In R. P. V. Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements* (p. 99 - 124). Oxford: Elsevier.
- Nodine, C., Kundel, H., Toto, L., & Krupinski, E. (1992). Recording and analyzing eye-position data using a microcomputer workstation. *Behavior Research Methods, Instruments, & Computers*, 24(3), 475–485.
- Noton, D., & Stark, L. (1971a). Eye movements and visual perception. *Scientific American*, 224(6), 34–43.
- Noton, D., & Stark, L. (1971b). Scanpaths in eye movements during pattern perception. *Science*, 171(3968), 308–311.
- Noton, D., & Stark, L. (1971c). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, 11(9), 929–942.
- Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*, 42(1), 188-204.
- Olsen, A. (2012). *The tobii i-vt fixation filter* (Tech. Rep.). Tobii Technology.
- Olsson, P. (2007). *Real-time and offline filters for eye tracking* (Unpublished master's thesis). KTH Royal Institute of Technology.
- Over, E., Hooge, I., Vlaskamp, B., & Erkelens, C. (2007). Coarse-to-fine eye movement strategy in visual search. *Vision Research*, 47(17), 2272 - 2280. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0042698907002088> doi: <http://dx.doi.org/10.1016/j.visres.2007.05.002>
- Paulsen, D. J., Hallquist, M. N., Geier, C. F., & Luna, B. (2015). Effects of incentives, age, and behavior on brain activation during inhibitory control: A longitudinal fmri study. *Developmental Cognitive*

- Neuroscience*, 11, 105 - 115. (Proceedings from the inaugural Flux Congress; towards an integrative developmental cognitive neuroscience)
- Peters, R. J., & Itti, L. (2008). Applying computational tools to predict gaze direction in interactive visual environments. *ACM Transactions on Applied Perception*, 5(2).
- Poulton, E. C. (1962). Peripheral vision, refractoriness and eye movements in fast oral reading. *British Journal of Psychology*, 53(4), 409–419.
- Privitera, C. M. (2006). The scanpath theory: its definition and later developments. In *Human vision and electronic imaging xi* (Vol. 6057, pp. 87–91).
- Privitera, C. M., & Stark, L. W. (2000). Algorithms for defining visual regions-of-interest: Comparison with eye fixations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(9), 970–982.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- Reimer, B., & Sodhi, M. (2006). Detecting eye movements in dynamic environments. *Behavior Research Methods*, 38(4), 667–682.
- Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review*, 101, 900–926.
- Rigas, I., Economou, G., & Fotopoulos, S. (2012). Biometric identification based on the eye movements and graph matching techniques. *Pattern Recognition Letters*, 33(6), 786–792.
- Rolfs, M. (2009). Microsaccades: Small steps on a long way. *Vision Research*, 49, 2415–2441.
- Rothkopf, C. A., & Pelz, J. B. (2004). Head movement estimation for wearable eye tracker. In *Proceedings of the eye tracking research & application symposium etra 2004*.
- Russo, J. E. (2010). Eye fixations as a process trace. In M. Schulte-Mecklenbeck, A. Kühlberger, & R. Ranyard (Eds.), *A handbook of process tracing methods for decision research* (pp. 43–64). Taylor & Francis.
- Saez de Urabain, I. R., Johnson, M. H., & Smith, T. J. (2015). Grafix: A semiautomatic approach for parsing low- and high-quality eye-tracking data. *Behavior Research Methods*, 47(1), 53–72. Retrieved from <http://dx.doi.org/10.3758/s13428-014-0456-0> doi: 10.3758/s13428-014-0456-0
- Salvucci, D. D., & Anderson, J. R. (1998). Tracing eye movement protocols with cognitive process models. In *Proceedings of the twentieth annual conference of the cognitive science society* (pp. 923–928).
- Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Etra '00: Proceedings of the 2000 symposium on eye tracking research & applications*.
- Santella, A., & DeCarlo, D. (2004). Robust clustering of eye movement recordings for quantification of visual interest. In *Etra '04: Proceedings of the 2004 symposium on eye tracking research & applications* (pp. 27–34). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/968363.968368> doi: 10.1145/968363.968368
- Santini, T., Fuhl, W., Kübler, & Kasneci, E. (2016). Bayesian identification of fixations, saccades, and smooth pursuits. In *Acm symposium on eye tracking research & applications, etra 2016*.
- Sauter, D., Martin, B. J., Di Renzo, N., & Vomscheid, C. (1991). Analysis of eye tracking movements using innovations generated by a kalman filter. *Medical and Biological Engineering and Computing*, 29(1), 63–69.
- Schwartz, S. H. (2010). *Visual perception* (fourth ed.; J. Morita & P. J. Boyle, Eds.). McGraw-Hill.
- Schwartz, S. H. (2013). *Geometric and visual optics* (second ed.). McGraw-Hill education.
- Scinto, L. F. M., & Barnette, B. D. (1986). An algorithm for determining clusters, pairs or singletons in eye-movement scan-path records. *Behavior Research Methods, Instruments, & Computers*, 18(1), 41–44.
- seeingmachines. (2005). *facelab 4 user manual* (4th ed.) [Computer software manual].
- Shellhamer, M. (1998). Nonlinear dynamic systems evaluation of 'rhythmic' eye movements (optokinetic nystagmus). *Journal of Neuroscience Methods*, 83(1), 45 - 56.
- Shellhamer, M., & Zalewski, S. (2001). A new application for time-delay reconstruction: detection of fast-phase eye movements. *Physics Letters A*, 291(4 - 5), 349 - 354.
- Sheth, B. R., & Young, R. (2016). Two visual pathways in primates based on sampling of space: Exploitation and exploration of visual information. *Frontiers in Integrative Neuroscience*, 10, 37. Retrieved from <http://journal.frontiersin.org/article/10.3389/fnint.2016.00037> doi: 10.3389/fnint.2016.00037
- Shic, F., Scassellati, B., & Chawarska, K. (2008). The incomplete fixation measure. In *Etra '08: Proceedings of the 2008 symposium on eye tracking research & applications* (pp. 111–114). New York, NY, USA: ACM.
- Smeets, J. B. J., & Hooge, I. T. C. (2003). Nature of variability in saccades. *Journal of Neurophysiol*, 90,

- 12–20.
- Smith, N. D., Crabb, D. P., Glen, F. C., Burton, R., & Garway-Heath, D. F. (2012). Eye movements in patients with glaucoma when viewing images of everyday scenes. *Seeing and Perceiving*, 25(5), 471–492.
- Smith, T. J., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. *Journal of Vision*, 13(8), 1–24.
- Smyrnis, N. (2008). Metric issues in the study of eye movements in psychiatry. *Brain and Cognition*, 68(3), 341–358.
- Stampe, D. M. (1993). Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behavior Research Methods, Instruments, & Computers*, 25(2), 137–142.
- Stark, L., & Ellis, S. R. (1981). Scanpath revisited: Cognitive models direct active looking. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: cognition and visual perception* (pp. 193–226). Hillsdale, NJ.
- Sundstedt, V., Stavrakis, E., Wimmer, M., & Reinhard, E. (2008). A psychophysical study of fixation behavior in a computer game. In *Apgv '08: Proceedings of the 5th symposium on applied perception in graphics and visualization* (pp. 43–50). ACM.
- Tafaj, E., Kasneci, G., Rosenstiel, W., & Bogdan, M. (2012). Bayesian online clustering of eye movement data. In *Proceedings of the symposium on eye tracking research and applications* (pp. 285–288). New York, NY, USA: ACM.
- Tatler, B., Wade, N., & Kaulard, K. (2007). Examining art: dissociating pattern and perceptual influences on oculomotor behaviour. *Spatial Vision*, 21(1), 165–184.
- Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, 7(14), 4. Retrieved from <http://dx.doi.org/10.1167/7.14.4> doi: 10.1167/7.14.4
- Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., & Velichkovsky, B. M. (2010). Yarbus, eye movements, and vision. *i-Perception*, 1, 7–27.
- Thornton, T. L., & Gilden, D. L. (2007). Parallel and serial processes in visual search. *Psychological Review*, 114. doi: 10.1037/0033-295X.114.1.71
- Tobii. (2014). Tobii studio version 3.3.0 (3.3.0 ed.) [Computer software manual].
- Tole, J. R., & Young, L. R. (1981). Digital filters for saccade and fixation detection. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception*. Lawrence Erlbaum.
- Trukenbrod, H. A., & Engbert, R. (2012). Eye movements in a sequential scanning task: Evidence for distributed processing. *Journal of Vision*, 12((1):5), 1–12.
- Tseng, P.-H., Carmi, R., Cameron, I. G. M., Munoz, D. P., & Itti, L. (2009). Quantifying center bias of observers in free viewing of dynamic natural scenes. *Journal of Vision*, 9(7), 4. Retrieved from <http://dx.doi.org/10.1167/9.7.4> doi: 10.1167/9.7.4
- Ungerleider, L. G., & Haxby, J. V. (1994). *ŠwhatŠ and ŠwhereŠ* in the human brain. *Current Opinion in Neurobiology*, 4(2), 157–165. Retrieved from <http://www.sciencedirect.com/science/article/pii/0959438894900663> doi: [http://dx.doi.org/10.1016/0959-4388\(94\)90066-3](http://dx.doi.org/10.1016/0959-4388(94)90066-3)
- Urruty, T., Lew, S., & Ihadaddene, N. (2007). Detecting eye fixations by projection clustering. *ACM Transactions on Multimedia Computing, Communications and Applications*, 3(4).
- Špakov, O., & Miniotas, D. (2007). Application of clustering algorithms in eye gaze visualizations. *Information Technology and Control*, 36(2), 213–216.
- Valsecchi, M., Gegenfurtner, K. R., & Schütz, A. C. (2013). Saccadic and smooth-pursuit eye movements during reading of drifting texts. *Journal of Vision*, 13(10), 8.
- van der Lans, R., Wedel, M., & Pieters, R. (2011). Defining eye-fixation sequences across individuals and tasks: the binocular-individual threshold (bit) algorithm. *Behavior Research Methods*, 43, 239–257.
- van der Linde, I., Rajashekar, U., Bovik, A. C., & Cormack, L. K. (2009). Doves: A database of visual eye movements. *Spatial Vision*, 22(2), 161–177. Retrieved from <http://live.ece.utexas.edu/research/doves>
- Van der Stigchel, S., Meeter, M., & Theeuwes, J. (2006). Eye movement trajectories and what they tell us. *Neuroscience and Biobehavioral Reviews*, 30, 666–679.
- van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1394), 359–366. Retrieved from <http://rspb.royalsocietypublishing.org/content/265/1394/359> doi: 10.1098/rspb.1998.0303
- Velichkovsky, B. M., Joos, M., Helmert, J. R., & Pan-

- nash, S. (2005). Two visual systems and their eye movements: evidence from static and dynamic scene perception. In *Proceedings of the xxvii conference of the cognitive science society* (p. 2283-2288). Retrieved from <https://tu-dresden.de/mn/psychologie/applied-cognition/ressourcen/dateien/publikationen/pdf/velichkovsky2005.pdf?lang=de>
- Vella, F., Infantino, I., & Scardino, G. (2016). Person identification through entropy oriented mean shift clustering of human gaze patterns. *Multimedia Tools and Applications*, 1–25.
- Veneri, G. (2013). Pattern recognition on human vision. In (p. 19-47). Transworld Research Network.
- Veneri, G., Piu, P., Federighi, P., Rosini, F., Federico, A., & Rufa, A. (2010, June). Eye fixations identification based on statistical analysis - case study. In *Cognitive information processing (cip), 2010 2nd international workshop on* (pp. 446–451).
- Veneri, G., Piu, P., Rosini, F., Federighi, P., Federico, A., & Rufa, A. (2011). Automatic eye fixations identification based on analysis of variance and covariance. *Pattern Recognition Letters*, 32, 1588–1593.
- Vidal, M., Bulling, A., & Gellersen, H. (2012). Detection of smooth pursuits using eye movement shape features. In *Proceedings of the symposium on eye tracking research and applications* (pp. 177–180).
- Špakov, O. (2012). Comparison of eye movement filters used in hci. In *Proceedings of the symposium on eye tracking research and applications* (pp. 281–284). ACM.
- Wass, S., Smith, T., & Johnson, M. (2013). Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behavior Research Methods*, 45(1), 229–250.
- Weiskrantz, L. (1972). Review lecture: Behavioural analysis of the monkey's visual nervous system. *Proceedings of the Royal Society of London B: Biological Sciences*, 182(1069), 427–455. Retrieved from <http://rspb.royalsocietypublishing.org/content/182/1069/427> doi: 10.1098/rspb.1972.0087
- Widdel, H. (1984). Operational problems in analysing eye movements. In *Theoretical and applied aspects of eye movement research* (Vol. 22, pp. 21–29).
- Wooding, D. S. (2002a). Eye movements of large populations: Ii. deriving regions of interest, coverage, and similarity using fixation maps. *Behavior Research Methods, Instruments, & Computers*, 34(4), 518–528.
- Wooding, D. S. (2002b). Fixation maps: quantifying eye-movement traces. In *Etra '02: Proceedings of the 2002 symposium on eye tracking research & applications* (pp. 31–36). New York, NY, USA: ACM.
- Wooding, D. S., Mugglestone, M. D., Purdy, K. J., & Gale, A. G. (2002). Eye movements of large populations: I. implementation and performance of an autonomous public eye tracker. *Behavior Research Methods, Instruments, & Computers*, 34(4), 509–517.
- Wyatt, H. J. (1998). Detecting saccades with jerk. *Vision Research*, 38, 2147–2153.
- Zangemeister, W. H., Stiehl, H. S., & Freksa, C. (Eds.). (1996). *Visual attention and cognition* (Vol. 116). North-Holland.
- Zemblys, R. (2016). Eye-movement event detection meets machine learning. In *The 20th international conference biomedical engineering 2016*. Retrieved from https://www.researchgate.net/publication/311027097_Eye-movement_event_detection_meets_machine_learning
- Zemblys, R., Niehorster, D. C., Komogortsev, O., & Holmqvist, K. (2017). Using machine learning to detect events in eye-tracking data (accepted paper). *Behavior Research Methods*.